

UNIVERSITÉ DU QUÉBEC

MÉMOIRE PRÉSENTÉ À
L'UNIVERSITÉ DU QUÉBEC À TROIS-RIVIÈRES

COMME EXIGENCE PARTIELLE
DE LA MAÎTRISE EN BIOLOGIE CELLULAIRE ET MOLÉCULAIRE

PAR
JESSICA MOISAN

IMPLICATIONS DES RÉSIDUS IONISABLES DANS LE CONTRÔLE DE
LA THERMOSTABILITÉ DE LA CARBOXYLESTÉRASE ESTGTA2

MAI 2017

Université du Québec à Trois-Rivières

Service de la bibliothèque

Avertissement

L'auteur de ce mémoire ou de cette thèse a autorisé l'Université du Québec à Trois-Rivières à diffuser, à des fins non lucratives, une copie de son mémoire ou de sa thèse.

Cette diffusion n'entraîne pas une renonciation de la part de l'auteur à ses droits de propriété intellectuelle, incluant le droit d'auteur, sur ce mémoire ou cette thèse. Notamment, la reproduction ou la publication de la totalité ou d'une partie importante de ce mémoire ou de cette thèse requiert son autorisation.

*Pour ma mère, sans qui je ne serais pas là où
j'en suis aujourd'hui.*

REMERCIEMENTS

Je voudrais d'abord remercier mon directeur de recherche, le professeur Marc Beauregard, qui m'a pris sous son aile dans le laboratoire dès le début de mes études au baccalauréat et tout au long de mes études supérieures et qui a été un directeur de recherche exemplaire. Il a su remplir sa mission et même en faire plus qu'on ne pouvait le demander. J'aimerais tout autant remercier ma collègue, Dre Fatma Meddeb-Mouelhi, qui m'a appris beaucoup de choses, tant au niveau des techniques de laboratoire, de la microbiologie, de la biologie moléculaire que de la vie en générale. J'ai eu la chance de vivre des moments inoubliables en leur compagnie et je les remercie du fond du cœur. J'aimerais aussi remercier tous ces nombreux collègues que j'ai eu la chance de côtoyer au laboratoire et avec qui j'ai pu échanger sur de nombreux sujets : Patrik Quessy, David Charbonneau, François Laframboise, Yannick Hébert-Ouellet, René Diene, Suzanne Attiori, Li Cui, Daniel Steven Agudelo, Nikolas Beauchesne, Vincent Bolduc, Guillaume L. Lemieux, Kevin Otis ainsi que Valérie Lalande. Je souhaite remercier particulièrement ma mère Nicole, ma sœur Fransheska, mon frère Samuel ainsi que mon conjoint Marc-Olivier qui m'ont appuyée tout au long de mon cheminement universitaire. Je tiens à remercier le regroupement québécois de recherche sur la fonction, la structure et l'ingénierie des protéines (PROTEO), le Conseil de recherches en sciences naturelles et en génie du Canada (CRSNG), le Décanat des études et la Fondation de l'Université du Québec à Trois-Rivières qui m'ont aussi appuyée pendant mes études. Je remercie aussi le Dr Albert Berghuis et Jonathan Blanchet de l'Université McGill, pour une collaboration de recherche en cristallographie. Je terminerai en remerciant mesdames Gloria Lavoie et Catarina Leote Franco Pio pour leur dévouement extraordinaire envers les étudiants et leur travail merveilleusement accompli ainsi que tout le personnel de l'Université du Québec à Trois-Rivières que j'ai côtoyé durant mes études.

AVANT-PROPOS

Ce mémoire fait l'objet du travail réalisé dans le cadre du programme de maîtrise en biologie cellulaire et moléculaire ainsi que d'un article scientifique soumis à un journal où il sera révisé par des pairs. Les travaux de recherche ont été réalisés entre 2014 et 2015 sous la supervision du professeur Marc Beauregard à l'Université du Québec à Trois-Rivières dans le laboratoire de biochimie des protéines et d'enzymologie industrielle.

L'enzymologie industrielle représente un secteur important de la biotechnologie. Depuis 2000, le marché mondial des enzymes industrielles a connu une importante croissance. Son chiffre d'affaires, qui était évalué à 3 milliards de dollars US en 2009, devrait atteindre 6 milliards d'ici l'an prochain. Les enzymes provenant de souches bactériennes thermophiles sont d'autant plus d'intérêt pour les industries puisqu'elles sont adaptées pour résister aux hautes températures qui sont couramment utilisées dans les procédés industriels (les réactions y sont accélérées et les réactifs mieux solubilisés). Comme la température des bassins de réaction est sujette à des fluctuations, il est nécessaire d'identifier des enzymes thermostables qui seront résistantes à ces fluctuations de température. Bien que plusieurs enzymes microbiennes thermostables aient été identifiées pour leur utilisation dans les procédés industriels, plusieurs de ces enzymes ne résistent pas aux conditions industrielles imposées (forces ioniques élevées, variations de pH). La caractérisation de nouvelles enzymes thermophiles est donc primordiale pour le développement d'enzymes industrielles adaptées aux procédés désirés. Dans cette perspective, la compréhension des relations entre la séquence d'acides aminés et les propriétés enzymatiques de ces enzymes est primordiale pour l'utilisation, le recyclage et la mise au point de variantes d'enzymes dans une application industrielle donnée. Cette compréhension permettrait de faire des prédictions (activité, stabilité et applications potentielles) lors de l'annotation génomique des séquences d'enzymes nouvellement identifiées, et permettrait de générer des variantes d'enzymes ayant des propriétés désirables pour une application donnée.

L'objectif principal de mon projet était d'étudier l'impact de certains facteurs environnementaux, mimant les conditions industrielles auxquelles l'enzyme pourrait être exposée, sur les propriétés physico-chimiques de la carboxylestérase EstGtA2 (Estérase provenant de *Geobacillus thermodenitrificans*, souche CMB-A2). Les résultats issus de ces expériences ont été présentés dans le cadre d'un séminaire de maîtrise, le 4 septembre 2015. Parallèlement, des expériences ont été réalisées en collaboration avec le Dr Albert Berghuis de l'Université McGill dans le but d'élucider la structure tridimensionnelle d'EstGtA2. L'élucidation de la structure tridimensionnelle, en plus des résultats présentés dans ce travail, permettra de valider le rôle et l'emplacement des acides aminés ionisés ciblés par ce travail, permettra d'orienter les mutations à apporter à l'enzyme pour augmenter sa stabilité dans différentes conditions et aidera à mieux comprendre la relation entre la séquence en acides aminés et la fonction de l'enzyme.

RÉSUMÉ

Parmi les enzymes ayant un potentiel d'application industrielle se trouvent les carboxylestérases (EC 3.1.1.1). Ces enzymes, qui sont indispensables dans la bioconversion des graisses dans la nature, catalysent l'hydrolyse des triacylglycérols. Les carboxylestérases agissent à pH alcalin et certaines sont actives et stables à haute température. Ce sont des enzymes présentant une gamme de propriétés physico-chimiques dignes du plus grand intérêt de la part des industriels dus à leurs propriétés enzymatiques diversifiées (spécificité du substrat, tolérance dans les solvants organiques). Ces enzymes trouvent des applications dans les industries du textile, de l'alimentaire, des pâtes et papiers, pharmaceutiques et cosmétiques. Par exemple, l'industrie des pâtes et papiers connaît de sérieux problèmes avec les extractibles lipidiques. Ces composés lipidiques, qui se retrouvent naturellement dans le bois des résineux, forment des particules colloïdales qui s'agrègent et se déposent sur les fibres de papiers, en plus de nuire à l'impression et à la qualité du papier. Par conséquent, les carboxylestérases pourraient résoudre ce problème dans l'industrie des pâtes et papiers.

Au cours de mon projet de maîtrise, la carboxylestérase EstGtA2 de *Geobacillus thermodenitrificans* a été caractérisée dans différentes conditions mimant les conditions industrielles. Mon but ultime étant d'identifier des éléments de structure clés pour le contrôle des propriétés d'EstGtA2 et ainsi contribuer à la compréhension fondamentale de son fonctionnement, la caractérisation de ces enzymes a été réalisée par des tests d'activité enzymatique sous différentes conditions de température, pH et de force ionique dans le but de pouvoir identifier le(s) pont(s) ionique(s) qui a(ont) un impact majeur sur les propriétés enzymatiques d'EstGtA2. Une caractérisation biophysique par des tests de dichroïsme circulaire a aussi été faite, où la dénaturation thermique des structures secondaires et tertiaires a été étudiée sous différents pH et différentes conditions de force ionique. Finalement, des mesures par diffusion dynamique de la lumière ont permis de comprendre le type d'agrégation et d'association subi par l'enzyme. Les résultats obtenus ont permis d'évaluer la capacité de l'enzyme à être recyclée après chauffage, de cibler certains acides aminés responsables du comportement de l'enzyme lors de dénaturation thermique dans différentes conditions et d'établir les différents chemins de dénaturations pris par l'enzyme dans ces conditions. Ces travaux permettent de mieux comprendre le comportement de l'enzyme lors de dénaturation thermique face à différentes conditions de pH et de force ionique. Ils introduisent une perspective pour la mutation dirigée dans le but de contrôler le type d'agrégation obtenue. En parallèle, des expériences de cristallographie de rayons X ont été réalisées en collaboration avec le Dr Albert Berghuis de l'Université McGill dans le but d'élucider la structure tridimensionnelle d'EstGtA2. À ce jour, aucun criblage n'a permis d'obtenir des cristaux d'assez bonne qualité pour obtenir une structure tridimensionnelle. Cependant, il a été possible de déterminer les conditions favorisant la formation de cristaux, instables dans le temps, pour cette enzyme. Un criblage plus exhaustif, avec des mutants plus stables d'EstGtA2, pourrait permettre d'obtenir des cristaux de meilleure qualité.

Mots-clés : Enzymes lipolytiques, estérase, thermophile, agrégation, interactions ioniques.

TABLE DES MATIÈRES

REMERCIEMENTS	iii
AVANT-PROPOS	iv
RÉSUMÉ.....	vi
LISTE DES TABLEAUX	x
LISTE DES FIGURES	xi
LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES	xii
LISTE DES SYMBOLES	xiv
CHAPITRE I	
INTRODUCTION.....	1
1.1 Les enzymes dans le monde du vivant	1
1.1.1 Composition des protéines : les acides aminés	1
1.2 Organisation structurale des protéines.....	4
1.3 Les enzymes et leurs applications	5
1.3.1 Les enzymes lipolytiques.....	7
1.4 Stabilité des enzymes à haute température	11
1.5 Agrégation des enzymes	14
1.6 Prédiction de structures et d'états dénaturés.....	16
1.7 Problème du repliement inversé	18
1.8 Annotation génomique	20
1.9 EstGtA2	23
1.10 Objectif	24
CHAPITRE II	
LE RÔLE DES RÉSIDUES CHARGÉS EN FONCTION DU PH DANS LA DÉNATURATION THERMIQUE DE LA CARBOXYLESTÉRASE ESTGTA2	25
2.1 Contribution des auteurs.....	25
2.2 Résumé de l'article	25
2.3 Article scientifique	27
Summary.....	28

Introduction	29
Materials and methods.....	32
Strains, plasmids and reagents	32
Protein expression and purification	32
CD spectroscopy	33
Dynamic light scattering	34
Isoelectric focusing	34
Enzyme activity measurements	35
Activity recovery after heat stress	35
Model of EstGtA2 and mapping of surface electrostatic potential	36
Results	36
Conformational and thermodynamic analyses of secondary structures..	36
Conformational and thermodynamic analysis of tertiary structures	38
Analysis of protein aggregation by dynamic light scattering	39
Determination of isoelectric point	40
Effect of pH and salt concentration on EstGtA2 activity	41
Investigation of activity recovery after heat treatment	41
Discussion.....	42
Proposed model for heat induced EstGtA2 denaturation and aggregation.....	46
Implications for industrial applications	47
Conflict of interest	48
Acknowledgements	49
Contribution of the authors.....	49
Conclusion.....	49
References	50
Tables	53
Legends to figures	55
Figures	57
Supplementary data	63

CHAPITRE III	
CONCLUSION.....	69
3.1 Synthèse et interprétation des résultats obtenus (sommaire).....	71
3.1.1 Analyse conformationnelle des structures secondaires	71
3.1.1 Analyse conformationnelle des structures tertiaires	72
3.1.2 Analyse de l'agrégation d'EstGtA2	73
3.1.3 Impact du point isoélectrique d'EstGtA2	73
3.1.4 Effet du pH et de la force ionique sur l'activité enzymatique	74
3.1.5 Effet du pH et de la force ionique sur la thermostabilité	74
3.1.6 Sommaire des analyses	75
3.2 Limitations.....	76
3.3 Perspectives de recherche	76
3.4 Conclusion	77
RÉFÉRENCES BIBLIOGRAPHIQUES.....	79

LISTE DES TABLEAUX

Tableau		Page
1.1	Valeurs de pK _a des chaînes latérales ionisables.....	4
1.2	Enzymes bactériennes et leurs applications industrielles	6
1.3	Propension d'acides aminés dans le cœur hydrophobe et en surface	12

LISTE DES FIGURES

Figure		Page
1.1	Les 20 acides aminés standards	2
1.2	Liaison peptidique entre deux acides aminés.....	3
1.3	Niveaux d'organisation des protéines.....	5
1.4	Repliement tridimensionnel des α/β hydrolases	8
1.5	Types de réactions catalysées par les enzymes lipolytiques.....	9
1.6	Organisation du couvercle sur le site actif des lipases.....	9
1.7	Cinétique des enzymes lipolytiques.....	10
1.8	Facteurs structuraux déterminants dans la stabilité des protéines	12
1.9	Formation de ponts salins	14
1.10	Théorie des paysages énergétiques	18
1.11	Comparaison des séquences disponibles sur UniProtKB versus les structures tridimensionnelles élucidées déposées sur PDB	19
1.12	Alignement de séquences multiples montrant l'évolution d'un pont salin près du site actif d'EstGtA2.....	22

LISTE DES ABRÉVIATIONS, SIGLES ET ACRONYMES

3D	Tridimensionnelle
ADN	Acide désoxyribonucléique
Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartate
CD	Circular dichroism
D	Acide aspartique
DMSO	Dimethyl sulfoxide
DO	Densité optique
Gln	Glutamine
Glu	Glutamate
H	Histidine
His	Histidine
Ile	Isoleucine
IMAC	Immobilized metal affinity chromatography
IPTG	Isopropyl- β -D-1-thiogalactopyranoside
LB	Luria Bertani
Leu	Leucine
Lys	Lysine
Met	Méthionine
Ni-NTA	Nickel-nitriloacetic acid
OD	Optical density

PDB	Protein Data Bank
Phe	Phénylalanine
pI	Point isoélectrique
pKa	Constante d'ionisation
p-NP	Para-nitrophenol
Rh	Rayon hydrodynamique
S	Sérine
SDS-PAGE	Sodium dodecylsulfate-Polyacrylamide gel electrophoresis
Ser	Serine
Thr	Thréonine
T_m	Température de fusion des structures secondaires
T_m^3	Température de fusion des structures tertiaires
Trp	Tryptophane
Tyr	Tyrosine
U.S.	United-State
Val	Valine

LISTE DES SYMBOLES

Å	Angström
L	Liter
M	Molar concentration
m	Milli
ms	Milliseconde
mg	Milligram
mL	Milliliter
n	Nano
λ	Longueur d'onde
μ	Micro
μ s	Microseconde
μ g	Microgram
μ L	Microliter
$[\Theta]$	Molecular ellipticity

CHAPITRE I

INTRODUCTION

1.1 Les enzymes dans le monde du vivant

Les protéines sont des macromolécules essentielles à la vie, jouant un rôle physiologique important dans tous les systèmes biologiques. En plus de gérer la transmission de l'information contenue dans l'ADN, les protéines sont impliquées dans une multitude de processus biologiques. Par exemple, elles sont impliquées dans le transport de certaines molécules (comme l'hémoglobine qui transporte l'oxygène), elles jouent un rôle prépondérant dans l'immunité (les anticorps), dans le mouvement (réplication de l'ADN, contraction musculaire), dans la reconnaissance cellulaire et elles agissent comme catalyseurs dans les voies métaboliques (enzymes). Une des caractéristiques qui rendent les protéines si uniques est la manière dont elles interagissent avec les autres molécules. Ce comportement, qui leur vient directement de leur structure, qui elle découle de leur composition en acides aminés, fait l'objet de nombreuses études en sciences de la vie.

1.1.1 Composition des protéines : les acides aminés

Les acides aminés sont les unités qui composent les protéines. On dénombre 20 acides aminés naturels chez les eucaryotes (Figure 1.1), les α -acides aminés, qui sont tous composés d'un groupement amine ($-\text{NH}_2$), d'un groupement carboxylique ($-\text{COOH}$), d'un hydrogène (H) et d'une chaîne latérale variable (R). Selon la polarité de leur chaîne latérale, les acides aminés ont été classés en trois groupes : non polaire (glycine, alanine, valine, leucine, isoleucine, méthionine, proline, phénylalanine et tryptophane), polaire chargé (lysine, arginine, histidine, acide aspartique, acide glutamique, aspartate et glutamate) et polaire non-chargé (sérine, thréonine, asparagine et glutamine) (Voet, Voet et Pratt, 2008).

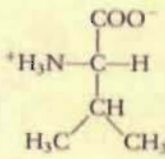
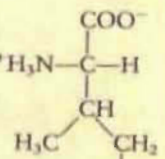
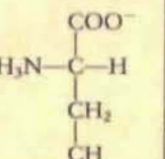
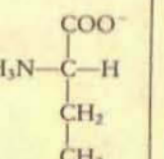
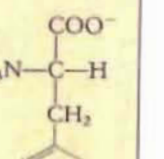
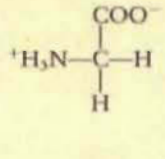
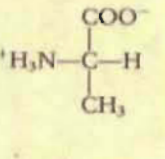
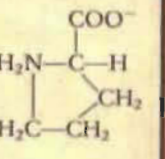
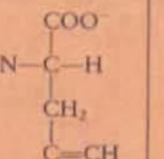
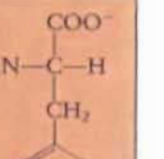
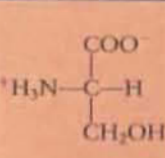
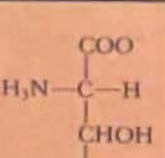
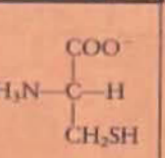
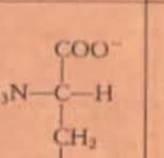
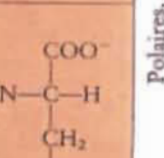
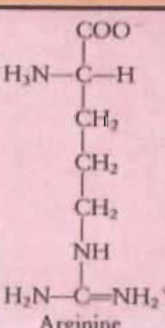
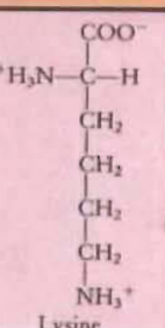
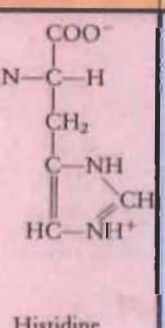
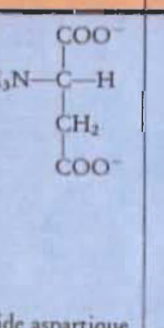
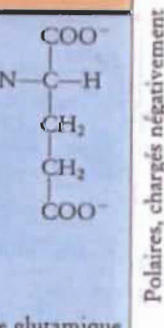
Non polaires	 Valine	 Isoleucine	 Leucine	 Méthionine	 Phénylalanine
	 Glycine	 Alanine	 Proline	 Tryptophane	 Tyrosine
	 Sérine	 Théonine	 Cystéine	 Asparagine	 Glutamine
	 Arginine	 Lysine	 Histidine	 Acide aspartique	 Acide glutamique
	Polaires, chargés positivement		Polaires, non-chargés		Polaires, chargés négativement

Figure 1.1 Les 20 acides aminés standards.

Les acides aminés sont généralement regroupés en trois classes (les non polaires, comprenant les aromatiques (haut), les polaires non-chargés (milieu) et les polaires chargés (bas de l'image). Source : Colin, Dictionnaire physique et biologie du XXI^e siècle. Document télé accessible à l'adresse <<http://www.gr-univers.fr/dictionnaire/index.php>>.

Dans les protéines, les acides aminés sont liés entre eux par des liaisons peptidiques, se formant au niveau du groupement amine et carboxylique (Figure 1.2). Une chaîne peptidique possède donc un résidu avec un groupement carboxylique (C-terminal) et un résidu avec un groupement amine (N-terminal) à ses extrémités. Les chaînes latérales interagissent de diverses façons avec le milieu environnant et ces interactions auront beaucoup d'influence sur le comportement de la protéine en milieu aqueux.

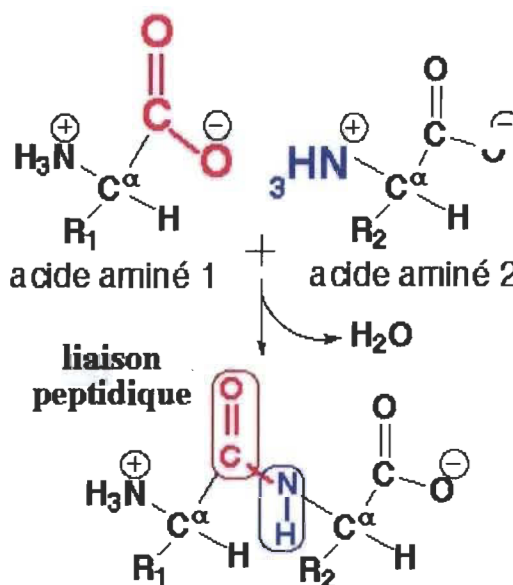


Figure 1.2 Liaison peptidique entre deux acides aminés.

La liaison peptidique se forme entre le groupement carboxylique (COO^- , représenté en rose) d'un acide aminé et le groupement amine (NH_3^+ , représenté en bleu) du second acide aminé. Les chaînes latérales (R_1 et R_2) sont attachées au carbone central (CH). Source : Éditions GR Univers, Les structures des protéines-Introduction. Document télé accessible à l'adresse < <http://biochimej.univ-angers.fr> >.

L'état d'ionisation des chaînes latérales exposées au solvant est tributaire du pH de la solution dans laquelle elles se trouvent ainsi que de leur microenvironnement, qui peut modifier de façon considérable le pK_a des chaînes latérales. Par exemple, une lysine enfouie dans une poche hydrophobe verra son pK_a perturbé jusqu'à une valeur de 7-8. La modification de leur état d'ionisation n'est seulement possible que pour les acides aminés qui présentent des groupements ionisables, pour lesquels des valeurs de constante d'ionisation (pK_a) théoriques ont été déterminées (Tableau 1.1). Les chaînes latérales des

acides aminés basiques (lysine, histidine et arginine) seront chargés positivement lorsque le pH est inférieur à leur pK_a , tandis que les acides aminés acides (acide glutamique et acide aspartique) seront chargés négativement lorsque le pH est supérieur à leur pK_a . En plus de ces derniers acides aminés, la cystéine et la tyrosine peuvent être sujettes à un changement d'état d'ionisation (leur groupement thiol ou alcool perdant un hydrogène) de leur chaîne latérale lorsqu'ils se trouvent à pH alcalin (chargés négativement lorsque le pH est supérieur à leur pK_a). Dans une protéine, les valeurs de pK_a dévient des valeurs théoriques, parfois même jusqu'à quelques unités de pH en fonction des chaînes latérales des résidus qui entourent la chaîne latérale en question.

Tableau 1.1

Valeurs de pK_a des chaînes latérales ionisables.
Source : Tiré de Voet, Voet et Pratt, 2008.

Acides Aminés	pK_a
Acide aspartique (aspartate)	3.90
Acide glutamique (glutamate)	4.07
Arginine	12.48
Cystéine	8.37
Histidine	6.04
Lysine	10.54
Tyrosine	10.46

1.2 Organisation structurale des protéines

L'organisation structurale des protéines a longtemps été questionnée et reste un sujet complexe. Ce n'est que dans les années 1950 que les premières structures tridimensionnelles eurent été résolues par diffraction de rayon X (tel que discuté dans Dickerson, 2008). À partir de ce moment, on comprit que les atomes de ces molécules étaient précisément organisés dans l'espace (Voet, Voet et Pratt, 2008). Le facteur le plus important dans la détermination de l'organisation d'une protéine est l'information

directement tirée de l'ADN, c'est-à-dire la séquence en acides aminés (Figure 1.3). L'enchaînement de ces derniers (la structure primaire) est décisif quant au repliement qui sera adopté par la protéine. La chaîne polypeptidique se replie et s'organise de manière à former des arrangements locaux périodiques appelés structures secondaires (hélices, feuillets, coudes et structures aléatoires). L'hélice alpha et le feuillet bêta sont reconnus comme étant les deux principales structures périodiques les plus rencontrées (Pauling, 1948, 1951, 1960) dans les protéines élucidées à l'époque (et jusqu'à maintenant). Le repliement des structures secondaires permettant de maximiser les interactions dues aux chaînes latérales et d'atteindre un haut niveau de compacité forme la structure tertiaire. Chez certaines protéines, un assemblage de plusieurs sous-unités tertiaires forme la structure quaternaire. Cette organisation structurale, qui confère de la stabilité aux protéines, leur confère une spécificité unique et détermine ainsi le rôle biologique de la protéine (Voet, Voet et Pratt, 2008).

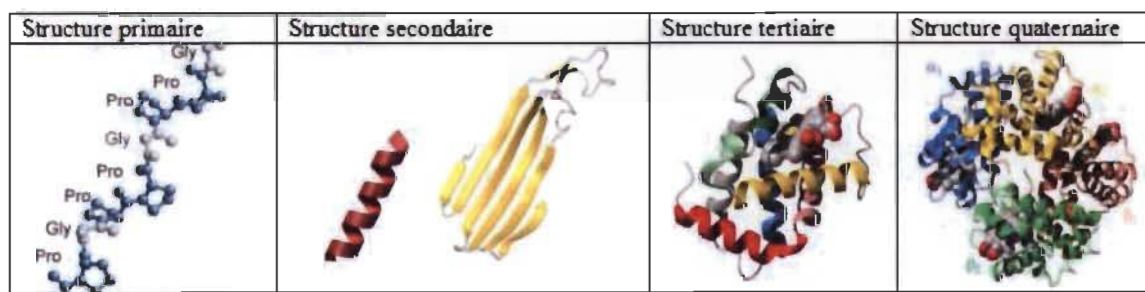


Figure 1.3 Niveaux d'organisation des protéines.

Les hélices (rouges) et les feuillets (jaunes) sont les structures secondaires les plus communes. Source : Lacouture, La Terre dans l'univers, la vie et l'évolution du vivant, Chapitre 13. Document télé accessible à l'adresse <<http://eric.lacouture.free.fr>>.

1.3 Les enzymes et leurs applications

La structure unique des enzymes leur permet une spécificité d'action unique avec les autres molécules. De ce fait, les enzymes agissent à titre de catalyseurs biologiques, permettant l'accélération de l'atteinte de l'équilibre d'une réaction chimique, en diminuant l'énergie d'activation. Comme les enzymes ont la capacité de former ou

de briser les liaisons chimiques, il n'est pas surprenant que les industries tentent de plus en plus de les inclure dans leurs procédés industriels. Depuis les années 2000, le marché mondial des enzymes industrielles ne cesse de croître. Étant estimé à 3 milliards de dollars US en 2009, il devrait dépasser les 6 milliards d'ici la fin de 2017 (Charnock et McCleary, 2005). L'utilisation des enzymes industrielles est reconnue pour ses avantages. En plus de diminuer le coût des matières premières et d'être rentable, la substitution des produits chimiques par les enzymes introduit une dimension écologique aux procédés industriels. L'utilisation des enzymes donne lieu à des procédés moins dommageables pour l'environnement en comparaison avec les procédés chimiques classiques. Ces molécules naturelles sont complètement dégradées dans l'environnement et leur utilisation a un impact important sur la réduction de consommation d'énergie des diverses industries. Plusieurs enzymes bactériennes (cellulases, xylanases, lipases/carboxylesterase) sont utilisées en industrie (Tableau 1.2) (Charnock et McCleary, 2005; Kirk, Borchet et Fuglsang, 2002). Grâce à leur capacité à hydrolyser les liens esters, les carboxylestérases trouvent des applications importantes en biotechnologie, que ce soit dans le textile, l'alimentation, la pharmacutique, en cosmétique et dans le secteur des pâtes et papiers.

Tableau 1.2

Enzymes bactériennes et leurs applications industrielles

Enzyme	Application	Secteur
Uréases	Élimination de l'urée	Vin
Xylanases	Élimination de la lignine	Papeterie
Hydrolases	Lyse des gels de biopolymères	Gaz et pétrole
Lipases	Dégradation des lipides	Détergents
Lipases	Modification de saveur	Alimentaire
Lipases	Transestérification	Pâtes et papiers

Par exemple, les lipases peuvent être utilisées dans la composition des détergents à lessive afin de dissoudre les graisses. Présentement, la plupart des détergents contiennent des produits néfastes pour l'environnement (phosphates, sulfates) et leur utilisation

contribue à l'augmentation du problème des algues bleues, par exemple (Bajpai et Tyagi, 2007). En utilisant les lipases, on peut contrôler la quantité de produits chimiques relâcher dans l'environnement et nettoyer nos vêtements de manière efficace et environnementale.

1.3.1 Les enzymes lipolytiques

Les enzymes lipolytiques (estérases et lipases) sont omniprésentes dans la nature (plantes, animaux, microorganismes) et sont reconnues pour leurs nombreuses utilisations dans les applications industrielles. Ces enzymes, qui catalysent les réactions d'hydrolyse et d'estérification, sont impliquées dans les voies métaboliques des lipides. Ces enzymes appartiennent à la famille des α/β hydrolases, de par leur structure formée de feuillets centraux entourés d'hélices (Figure 1.4) (Ollis *et al.*, 1992). Leur site actif est composé d'une triade catalytique d'acides aminés très conservés dans l'évolution (Ser-Asp/Glu-His). Les enzymes lipolytiques ont fait l'objet de plusieurs études dans les dernières années et sont toujours d'actualité (Charbonneau, Meddeb-Mouelhi et Beauregard, 2010; Shao, Xu et Yan, 2014). Ces enzymes ont initialement été classifiées en huit familles en fonction de motifs d'acides aminés conservés et de leurs propriétés biologiques par Arpigny et Jaeger (1999), classification qui ne cesse de s'accroître (Charbonneau et Beauregard, 2013). Ce classement enzymatique permet maintenant de faire des prédictions au niveau de facteurs structuraux lors de l'identification de nouvelles enzymes lipolytiques. D'un autre côté, ces enzymes ont aussi été classifiées en 38 familles homologues, comprenant 16 super familles, en fonction de leur séquence, de leur structure et de leur fonction par Fisher et Pleiss (2003).

Les carboxylestérases (E.C.3.1.1.1) et les lipases (E.C.3.1.1.3) catalysent l'hydrolyse et la synthèse des liens esters des molécules sur des substrats à longueur de chaînes variables (Figure 1.5). Leur spécificité est donc plus souple que chez la plupart des enzymes. Les estérases sont actives sur des substrats à courtes chaînes (< 10 carbones), contrairement aux lipases qui sont plutôt actives sur les substrats à longues chaînes (> 10 carbones) (Charbonneau et Beauregard 2013). Une autre caractéristique qui différencie les lipases des estérases est la présence d'un couvercle chez

les lipases couvrant le site actif (Figure 1.6). Comme leurs substrats sont moins solubles (longues chaînes aliphatiques) que ceux des estérases (courtes chaînes), les réactions catalysées par les lipases se produisent à l'aide d'un système d'activation inter-faciale, qui se fait à l'interface lipide-eau, là où entre en jeu le couvercle de l'enzyme. Lorsque la lipase s'approche d'une surface lipidique, le couvercle s'ouvre de manière à laisser pénétrer le substrat dans la cavité du site actif. Les carboxylestérases n'usent pas d'un tel mécanisme et n'ont pas à être activées. De ce fait, les carboxylestérases suivent généralement une cinétique enzymatique de type Michaelis-Menten (Figure 1.7) puisqu'elles n'ont pas besoin d'être activées par la présence d'un substrat tandis que les lipases suivent une cinétique à tendance sigmoïdale (la présence d'un substrat va activer l'ouverture du couvercle et induit un délai supplémentaire dans la cinétique) (Bornscheuer 2002; Fojan *et al.*, 2000; Kourist, Brundiek et Bornscheuer, 2010).

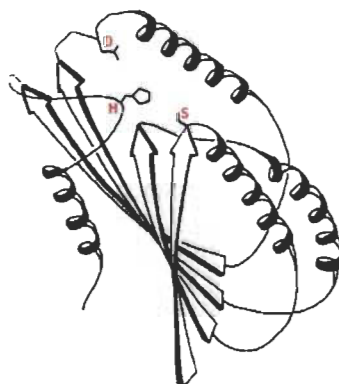


Figure 1.4 Repliement tridimensionnel des α/β hydrolases.

Les acides aminés conservés (Sérine (S), Acide aspartique (D) et Histidine (H)) qui sont impliqués dans la triade catalytique des enzymes de type α/β hydrolases sont montrés en rouge. Les feuillets sont représentés par des flèches et les hélices par des serpentins. Source : inspiré de Ollis *et al.*, 1992.

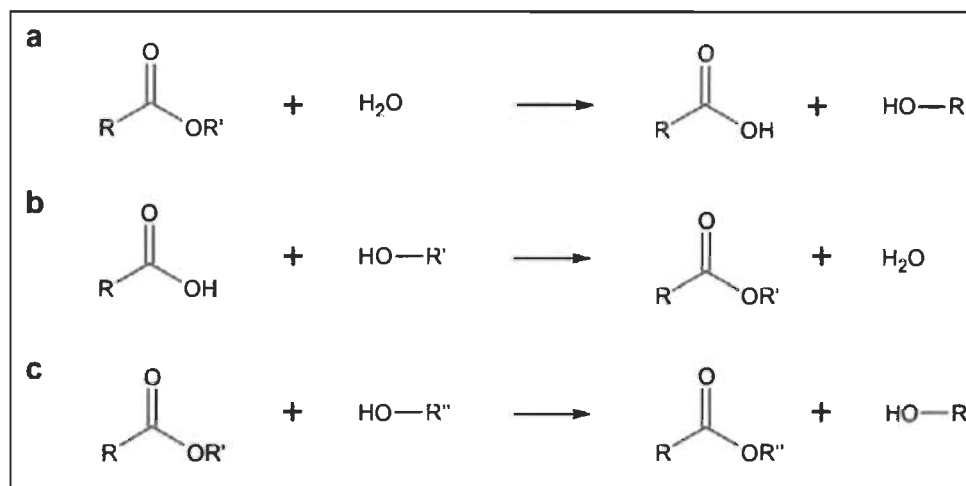


Figure 1.5 Types de réactions catalysées par les enzymes lipolytiques.

Les enzymes lipolytiques peuvent catalyser différentes réactions comme celle d'hydrolyse, nécessitant la présence de molécule d'eau (H_2O) (a), de synthèse, nécessitant la présence d'une molécule avec une chaîne aliphatique (b) et de transestérification, nécessitant aussi la présence d'une molécule avec une chaîne aliphatique (c). Source : Charbonneau, 2014.

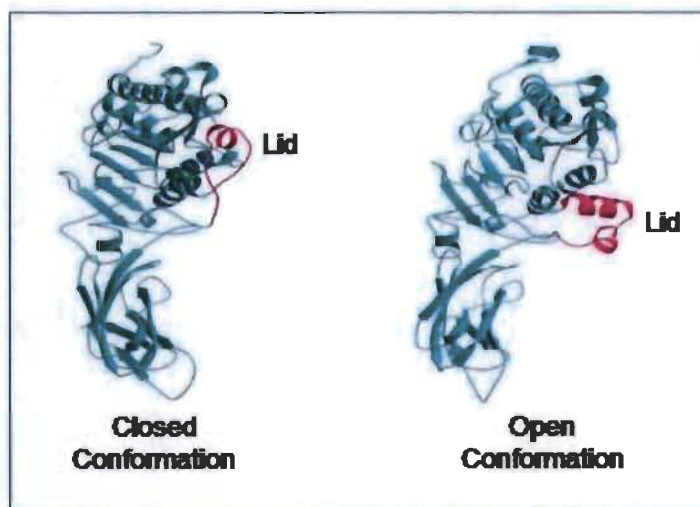


Figure 1.6 Organisation du couvercle sur le site actif des lipases.

Les lipases sont structurées de manière à avoir leur couvercle fermé (en mauve sur l'image de gauche) protégeant le site actif lorsqu'aucun substrat lipidique n'est à proximité (les acides aminés du site actif étant généralement plus hydrophobes que polaires, ils ne peuvent être exposés librement au milieu aqueux). Lorsqu'un substrat (un lipide) s'approche du site actif, l'affinité entre le lipide et les acides aminés hydrophobes fait que le couvercle s'ouvre de manière à faire entrer le substrat dans la cavité catalytique (à droite). Source : Palomo, 2003.

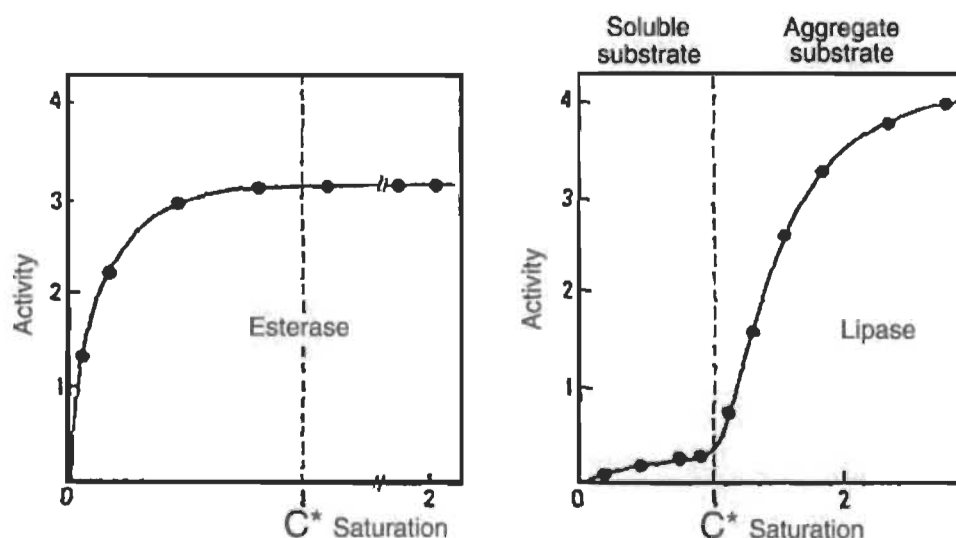


Figure 1.7 Cinétique des enzymes lipolytiques.

Les estérases (gauche) suivent une cinétique de type Michaelis-Menten (la cinétique n'est pas ralentie par l'ouverture du couvercle) tandis que les lipases (droite) doivent passer par l'activation inter-faciale et montrent une cinétique sigmoïdale. Source : Charbonneau, 2014.

Les carboxylestérases (estérases) possèdent plusieurs propriétés physico-chimiques intéressantes (spécificité, tolérance aux solvants organiques) d'un point de vue industriel et trouvent des applications dans plusieurs domaines (alimentation, cosmétique, pharmaceutique, fabrication de détergents, pâtes et papiers) (Bornscheuer 2002; Yu *et al.*, 2010). Les estérases provenant de souches extrêmophiles sont d'autant plus d'intérêt pour les procédés industriels, car elles sont adaptées pour résister et conserver leur activité dans des conditions extrêmes (pH, température, concentrations ioniques). Dans les bassins industriels impliquant des substrats liquides, la température élevée contribue à diminuer la viscosité du milieu et ainsi à augmenter la vitesse de conversion des substrats (vélocité) et la solubilité de ceux-ci (Van den Berg, 2003). Jusqu'à maintenant, beaucoup de carboxylestérases microbiennes ont été identifiées dans le but de les introduire dans des procédés industriels (Shao, Xu et Yan, 2014; Yu *et al.*, 2010). Malheureusement, beaucoup d'entre elles échouent, car elles ne sont pas adaptées pour résister à toutes les conditions extrêmes rencontrées (pH, température, concentrations ioniques) (Van den Berg, 2003; Li *et al.*, 2014). De ce fait, l'identification et la caractérisation de nouvelles estérases extrêmophiles est d'un des plus grands intérêts pour le développement

d'enzymes stables, et donc recyclables dans les procédés industriels. En plus d'apporter de nouvelles connaissances qui permettront de développer des enzymes mieux adaptées aux procédés industriels, les études de caractérisation sur les enzymes lipolytiques permettront d'augmenter la quantité d'information qui sera tirée directement des séquences d'ADN disponibles à mesure qu'on séquence de nouveaux génomes (annotation génomique). La caractérisation des relations séquence-structure-propriétés dans divers regroupements évolutifs d'enzymes permettra de mieux comprendre les signatures ou motifs de séquences qui déterminent les propriétés. Cette information cruciale pourra par la suite être utilisée dans la prédiction de propriétés enzymatiques à partir de la séquence d'ADN (annotation génomique). Dans l'ère actuelle de séquençage à haut débit où l'identification de nouveaux gènes se fait beaucoup plus rapidement que la caractérisation des enzymes encodées par ces gènes, l'interprétation d'information encodée directement à partir des séquences devient primordiale.

1.4 Stabilité des enzymes à haute température

De nos jours, les facteurs impliqués dans la stabilité conformationnelle des enzymes sont bien connus (Figure 1.8). L'effet hydrophobe constitue le facteur le plus déterminant pour le repliement tertiaire. Lors du repliement, les acides aminés non polaires (Val, Leu, Ile, Met et Phe) sont généralement regroupés au centre de l'enzyme, formant un cœur hydrophobe, de manière à minimiser les contacts entre ceux-ci et l'environnement aqueux dans lequel se trouve l'enzyme (Tableau 1.3). Les acides aminés polaires (Arg, His, Lys, Asp et Glu) sont normalement exposés à la surface, faisant contact avec le milieu aqueux. Les résidus polaires non-chargés (Ser, Thr, Asn, Gln et Tyr) sont aussi dispersés à la surface. Les interactions électrostatiques de type forces de van der Waals, les ponts hydrogène ainsi que les ponts ioniques (souvent appelés ponts salins) ne contribuent que très peu à la stabilité conformationnelle à température physiologique (des ponts hydrogènes se forment avec le solvant même lorsque l'enzyme est dépliée). Outre ces derniers, on discerne les ponts disulfures (formés entre deux cystéines lors de l'oxydation des groupements thiols) qui contribuent à la stabilité de la structure tertiaire (Voet, Voet et Pratt, 2008).

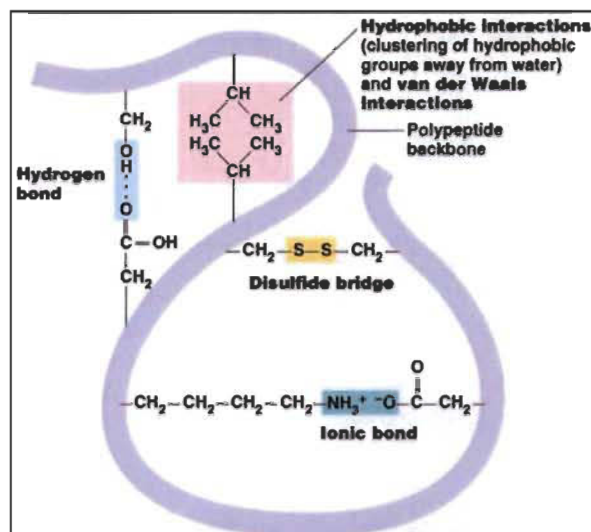


Figure 1.8 Facteurs structuraux déterminants dans la stabilité des protéines.

Les interactions entre les chaînes latérales sont déterminantes pour le contrôle de la stabilité conformationnelle. On retrouve notamment les liaisons hydrogènes (en bleu, haut à gauche), l'effet hydrophobe et les forces de van der Waals (en rose, haut à droite), les ponts disulfures (en jaune, au milieu) et les ponts ioniques (en bleu, en bas). Source : Charbonneau, 2014.

Les ponts salins sont des interactions ioniques formés entre deux résidus ionisables de charges opposées. Pour qu'une interaction se produise, les résidus doivent se trouver assez près l'un de l'autre ($< 4\text{\AA}$) et ils doivent se trouver dans une orientation favorable. La formation de ponts salins est aussi dépendante du pH de l'environnement. Pour qu'un pont puisse se former, il faut que les deux résidus (l'un acide et l'autre basique) soient ionisés au pH en question. Par exemple, pour qu'un pont salin puisse se former entre un glutamate et une lysine, on doit se trouver dans des conditions de pH variant entre 3.7 et 10.2 tandis qu'un pont salin entre un aspartate et une histidine peut se former entre pH 4.1 et 6.1 (Figure 1.9). Plusieurs études ont permis de reconnaître que les enzymes thermostables sont caractérisées par une présence accrue de ponts salins à leur surface. Ces ponts salins sont beaucoup plus stabilisants à haute température, où l'eau devient moins polaire. En plus de trouver beaucoup plus de ponts salins, ceux-ci sont organisés sous forme de réseaux dans les protéines stables à haute température (Kumar, Tsai et Nusinov, 2000; Vetriani *et al.*, 1998; Knochel *et al.*, 2002). En comparant les structures 3D d'enzymes mésophiles avec leurs homologues thermophiles, les experts ont pu

clairement démontrer le rôle de ces facteurs structuraux clés (les ponts salins) dans la stabilité thermique des enzymes. De plus, une augmentation du nombre de ponts hydrogènes formés entre les chaînes latérales, une augmentation du nombre d'arginine et de tyrosine, ainsi qu'une diminution du nombre de résidus thermolabiles (cystéine et sérine) ont été rapportés par Kumar, Tsai et Nussinov (2000). Comme ces derniers facteurs (ponts salins, ponts hydrogène entre chaînes latérales, modification de la présence de certains acides aminés) ont été observés chez différentes familles d'enzymes, on pense qu'ils sont des facteurs clés pour la stabilité des enzymes à haute température, comparativement à certains facteurs (meilleure compacité hydrophobe, diminution du nombre de prolines) qui ne sont conservés que dans certaines familles individuelles. Parmi les différentes études citées par Kumar, Tsai et Nussinov (2000), les ponts salins constituent et restent une cible de choix pour l'étude des facteurs contrôlant la stabilité des enzymes à haute température.

Tableau 1.3

Propension d'acides aminés dans le cœur hydrophobe et en surface

Acides aminés	Indice d'hydrophobicité	Acides aminés	Indice d'hydrophobicité
Isoleucine	4,5	Sérine	-0,8
Valine	4,2	Tyrosine	-1,3
Leucine	3,8	Proline	-1,6
Phénylalanine	2,8	Histidine	-3,2
Cystéine	2,5	Acide glutamique	-3,5
Méthionine	1,9	Glutamine	-3,5
Alanine	1,8	Acid aspartique	-3,5
Glycine	-0,4	Asparagine	-3,5
Thréonine	-0,7	Lysine	-3,9
Tryptophane	-0,9	Arginine	-4,5

L'index d'hydrophobicité (selon l'échelle de Kyle-Dolittle) représente la tendance des acides aminés à se positionner à la surface (+) ou dans le cœur hydrophobe (-). Source : Inspiré de Stipanuk. Document télé accessible à l'adresse <<https://basicmedicalkey.com>>.

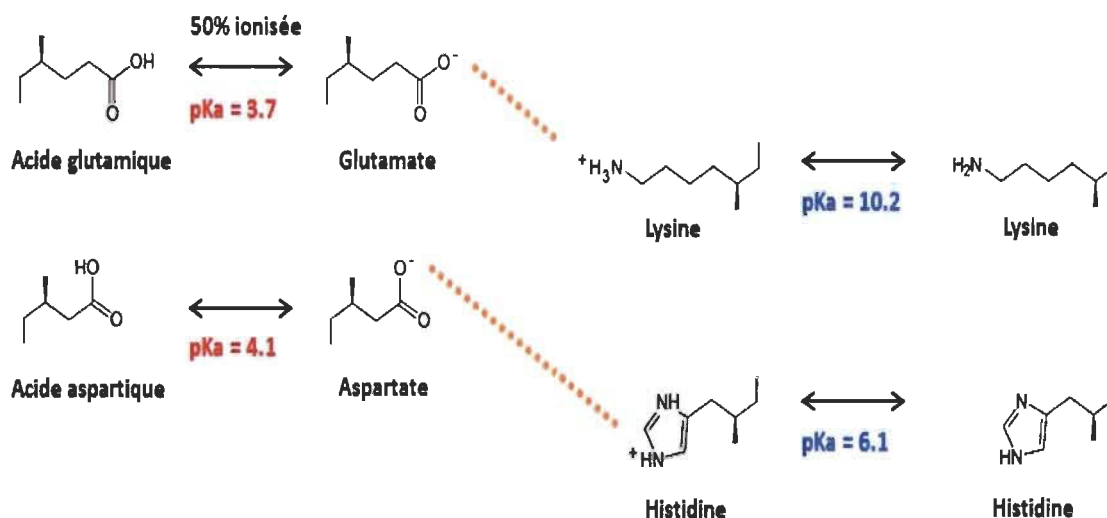


Figure 1.9 Formation de ponts salins.

La formation de ponts salins dépend du pH environnant. Un pont salin entre un glutamate et une lysine peut se former dans une plage de pH 3.7-10.2 (haut) tandis qu'un pont entre un aspartate et une histidine peut se former entre 4.1 et 6.1 (bas).

1.5 Agrégation des enzymes

L'ensemble des forces qui stabilisent les structures des enzymes les rendent aussi susceptibles à une dénaturation et à leur agrégation. L'agrégation est un phénomène où les enzymes se collent les unes aux autres de manière à former un complexe insoluble dans le milieu aqueux. Les enzymes sont sensibles aux conditions dans lesquelles elles se trouvent. La chaleur peut entraîner la fusion de l'enzyme et une variation de pH peut désorganiser les interactions ioniques nécessaires à la stabilité conformationnelle. Par exemple, lorsque le pH du milieu est près du pI (pH où la protéine est globalement neutre), les protéines vont avoir tendance à s'agréger. Leur repliement tridimensionnel ainsi que leur activité peuvent être modifiés en fonction des conditions environnantes.

L'agrégation des protéines est un problème de taille dans plusieurs domaines reliés à la biotechnologie. Au niveau de la mise en marché de vaccins et traitements thérapeutiques, la présence d'agrégats de protéines est un problème majeur considérant le risque important d'un agrégat face à des réactions immunitaires (Roberts, 2014).

Certains processus d'agrégation sont aussi impliqués dans de nombreuses maladies. Par exemple, les protéines formant des structures amyloïdes dans la maladie de l'Alzheimer (Fink, 1998; Tsolis *et al.*, 2012). Les enzymes qui s'agrègent ne sont plus fonctionnelles, ce qui diminue la rentabilité d'un processus industriel où elles sont impliquées. Il est donc impératif de trouver des moyens pour contrôler ces processus d'agrégation.

Que ce soit pour le développement de drogues thérapeutiques ou d'enzymes industrielles, la compréhension des différents mécanismes de dénaturation et d'agrégation permettra de mieux comprendre et prévenir l'agrégation des enzymes. Au niveau des applications industrielles, le contrôle de ces propriétés physiques permettra de mettre au point des enzymes résistantes pour les applications désirées et de diminuer le temps et les coûts reliés à leur utilisation.

À ce jour, les processus par lesquelles les enzymes s'agrègent restent hypothétiques. Des études récentes ont démontré que l'agrégation serait tributaire de trois phénomènes particuliers, en plus de la concentration en protéines : les conditions environnementales de l'échantillon, la structure tertiaire adoptée par l'enzyme et la séquence en acides aminés (Roberts, 2014).

L'agrégation irréversible des enzymes peut se manifester lorsque plusieurs unités se retrouvent dans une conformation non-native. La modification de la structure tertiaire peut être causée par un phénomène extérieur causant une distorsion de la structure tertiaire, par exemple, lorsque les enzymes sont soumises à des conditions extrêmes de température. La distorsion de la structure tertiaire entraîne l'exposition de certaines séquences peptidiques cibles, les « aggregation-prone sequences », qui sont responsables de causer l'agrégation. On dit de ces « hot spots » qu'ils sont souvent des séries d'acides aminés hydrophobes, ou faiblement chargés et qui sont sujets à former des feuilletts bêta ou des hélices alpha avec les brins à proximité (Tsolis *et al.*, 2012).

Des études mettent aussi de l'avant l'importance qui doit être accordée aux acides aminés chargés qui sont directement impliqués dans la stabilité conformationnelle et donc dans les processus d'agrégation (Tsai *et al.*, 2015). Habituellement, la formation du premier noyau d'agrégats entraîne les autres sous-unités à s'assembler de manière à former ces complexes stables (Fink, 1998; Roberts, 2014). La dénaturation des protéines ne donne pas toujours lieu à des espèces dénaturées agrégées dont l'état est irréversible. Une association contrôlée peut se présenter lorsque plusieurs sous-unités, plutôt bien repliées, s'auto-assemblent entre elles. Aussi, selon les conditions environnantes, c'est-à-dire une augmentation de la température de courte durée, les enzymes peuvent se dénaturer (se déplier) sans qu'il n'y ait d'association entre les différentes unités. Cet état, soluble et réversible, n'atteint pas un état d'agrégation macroscopique (Sassi *et al.*, 2011). Une fois que la température est abaissée à la normale, l'enzyme peut se replier et retrouver son activité enzymatique. Lorsqu'on parle de capacité de repliement et de restauration de l'activité résiduelle après chauffage, on parle de la thermostabilité. Attention de ne pas confondre la thermostabilité avec la stabilité thermique.

La stabilité thermique est caractérisée par une mesure de température à laquelle l'enzyme conserve 50 % de ses structures (ou autre propriété observable et reliée à la conformation). L'enzyme prendra le chemin vers une association soluble ou vers un état d'agrégation non soluble dépendamment des conditions dans lesquelles elle se trouve (Roberts, 2014). Les conditions environnementales sont donc déterminantes dans le type de stress (durée de dénaturation thermique, pH, force ionique) imposé à l'enzyme, aux séquences exposées ainsi qu'aux forces électrostatiques impliquées dans leur association. La morphologie des agrégats varie beaucoup pouvant être très ordonnée comme les amyloïdes ou pas ordonnée du tout, comme les corps d'inclusion (Roberts, 2014; Fink, 1998).

1.6 Prédiction de structures et d'états dénaturés

Des études récentes ont mis en évidence la conception d'outils de prédiction intéressants qui permettent de prédire les séquences en acides aminés impliqués et le type

de conformation adoptée dans les protéines. Toutefois, il faut savoir que ces algorithmes travaillent à partir de séquences protéiques linéaires. Il n'est donc pas évident de prédire avec précision quelles sections de séquence seront exposées en fonction des conditions expérimentales (Tsolis *et al.*, 2012).

Les recherches sur le comportement des enzymes en solutions deviennent donc cruciales pour l'optimisation de ses outils, pour la compréhension des mécanismes d'agrégation et leur prévention. Depuis longtemps, on essaie de comprendre ce qui guide les protéines dans leur repliement. Est-ce seulement la séquence qui guide ces processus de dénaturation? Quels sont les liens entre la séquence et le type d'agrégat qui sera formé? Ces questionnements ont fait naître plusieurs théories sur le processus de repliement des protéines. On se dit que les protéines doivent bien se dénaturer ou se replier selon un chemin préférentiel, selon un « paysage énergétique » (Figure 1.10) dépendant des conditions environnementales qui entourent l'enzyme en question (Onuchic *et al.*, 1994; Dill, 1997).

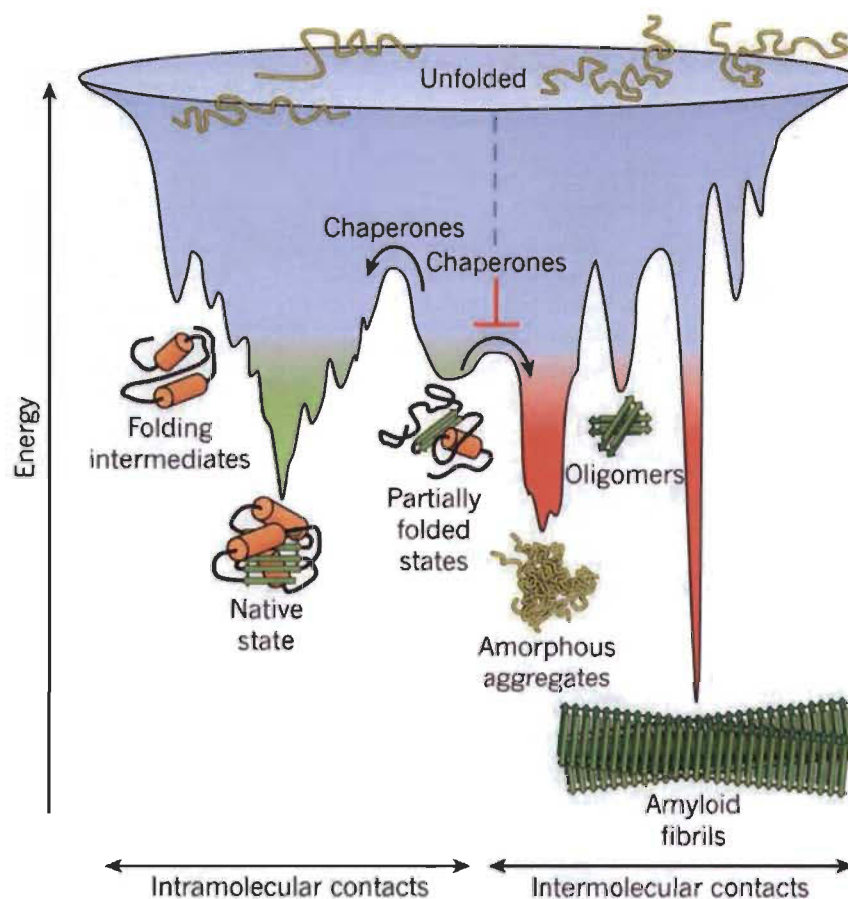


Figure 1.10 Théorie des paysages énergétiques.

Cette figure présente d'abord différents chemins énergétiques qui mènent à différents états énergétiques impliqués dans le repliement et la stabilité des protéines (à gauche; intermédiaires de repliement, état natif et état de repliement intermédiaire). Ensuite, on montre les états impliqués lors de la dénaturation, l'agrégation ou la cristallisation des protéines (à droite; agrégats amorphe, oligomère et fibrille de type amyloïde). Source : Harti *et al.*, 2011.

1.7 Problème du repliement inversé

Depuis l'arrivée des projets de séquençage à haut débit, le nombre de séquences protéiques disponibles dans les banques de données connaît une croissance exponentielle. En 2007, environ 5.3 millions de séquences étaient déposées dans la base de données

UniProtKB¹, et ce, sans compter les autres enzymes présentes dans la nature qui n'ont pas été identifiées. En revanche, le nombre de structures tridimensionnelles élucidées à la même époque connaissait un décalage important en raison des difficultés techniques, du temps et des coûts reliés (Figure 1.11). Pour la même année (2007), on compte environ 44 000 structures déposées sur PDB², ce qui représente moins de 1 % des séquences déposées dans UniProKb. Malheureusement, cet écart ne cesse d'augmenter (Ridgen, 2009). L'écart entre l'information tirée des séquences disponibles et celles tirées des structures tridimensionnelles cause un ralentissement des connaissances et du développement des enzymes dans plusieurs domaines, comme le développement de variantes d'enzymes pour les applications industrielles ou thérapeutiques.

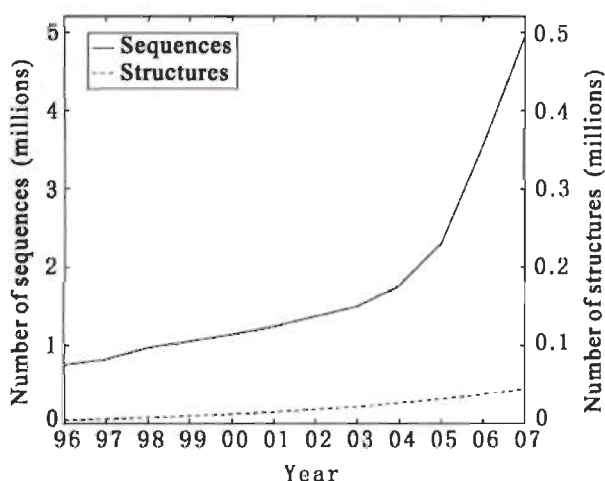


Figure 1.11 Comparaison des séquences disponibles sur UniProtKB versus les structures tridimensionnelles élucidées déposées sur PDB.

Le nombre de séquences disponibles (gauche) augmente de façon exponentielle tandis que le nombre de structures élucidées (droite) n'augmente presque pas pour les 12 dernières années. Source : Ridgen, 2009.

Pour obtenir de l'information sur la structure ou les propriétés enzymatiques (stabilité, comportement en solution) des enzymes nouvellement identifiées à partir de

¹ UniProtKB est une base de données dont la mission est de fournir gratuitement des informations sur les séquences et la fonction des protéines (<http://www.ebi.ac.uk/uniprot>).

² Protein Data Bank (PDB) est une banque de données accessible qui constitue un portail d'information sur les structures de protéines (<http://www.rcsb.org/pdb/home/home.do>).

leur séquence, il faut les produire et les caractériser en laboratoire, ce qui prend bien plus d'efforts et d'argent que le simple séquençage de son ADN. Le problème du repliement des protéines est toujours d'actualité. On ne sait toujours pas quelle est la relation précise qui existe entre l'information contenue dans la séquence en acides aminés (1D) et l'information démontrée par la structure 3D et de ce fait, il nous est impossible de prédire la structure tertiaire des enzymes avec leur séquence uniquement (Science Editorial, 2005). Bien qu'ils existent une panoplie de logiciels pour faire du « template based-modelling », c'est-à-dire modéliser la structure d'une protéine à partir de la structure d'une protéine homologue, ces techniques connaissent leurs défauts. Elles peuvent permettre d'obtenir un squelette fiable de la structure, mais cela ne nous aide pas à comprendre pourquoi l'enzyme adopte cette conformation. De plus, ces modélisations ne permettent pas de comprendre les effets sur la structure lors de mutations d'acides aminés précis. S'il n'y a pas de protéine homologue identifiée et élucidée, on peut tenter d'élucider la structure « *ab initio* », soit utiliser des logiciels de prédiction de structure qui se base sur un principe d'énergie. La modélisation, qui génère plusieurs repliements différents, proposera un modèle final pour lequel l'état sera le plus stable thermodynamiquement parlant et l'énergie de repliement le plus faible. Ici encore, les résultats sont à prendre avec un grain de sel (Ridgen, 2009). Rappelons que la protéine *in vivo* trouve son repliement optimal moins de 10 ms et parfois même de l'ordre des μ s lors de sa sortie du ribosome. Si une protéine de taille moyenne (150 acides aminés) tentait chaque conformation possible dans le but de trouver son état où l'énergie de repliement est le plus faible, elle devrait essayer 3^{300} conformations possibles, ce qui représenterait bien plus que des années. À ce jour, les mêmes questions se posent. Existe-t-il un code pour le repliement des protéines? Quels sont les mécanismes par lesquels une séquence d'acides aminés adopte sa structure finale aussi rapidement? Comment pourrait-on prédire le comportement d'une enzyme par sa séquence en acides aminés?

1.8 Annotation génomique

L'annotation génomique est une méthode de bio-informatique qui consiste à détecter des informations dans les données de séquençage. Cette méthode dépend de la

génération de ces informations, qui elle passe par l'expérimentation et par la comparaison d'enzymes homologues ou distantes. Par exemple, on a pu identifier des éléments structuraux stabilisants les enzymes à haute température en comparant des enzymes mésophiles avec leurs homologues thermophiles (réseaux de ponts salins discutés plus haut). Pour ce faire, on a recours à l'alignement multiple de séquences. En comparant des enzymes qui se comportent de la même façon, on peut repérer des motifs importants dans leur séquence. Concernant EstGtA2, en comparant des séquences de résidus d'acides aminés d'estérases (Figure 1.12), on a pu repérer un pont salin près du site actif qui est conservé dans l'évolution et même dans les organismes les plus distants (Charbonneau et Beauregard, 2013). Ceci est une véritable indication que le pont formé par ces résidus doit être d'une importance majeure. De plus, dans une sous-famille d'estérase, il y a une inversion de polarité parmi les résidus qui forment ce pont dans l'évolution, ce qui veut dire que quand la nature défait le pont, la nature l'a corrigé même si ça impliquait d'inverser la polarité de l'autre l'acide aminé.

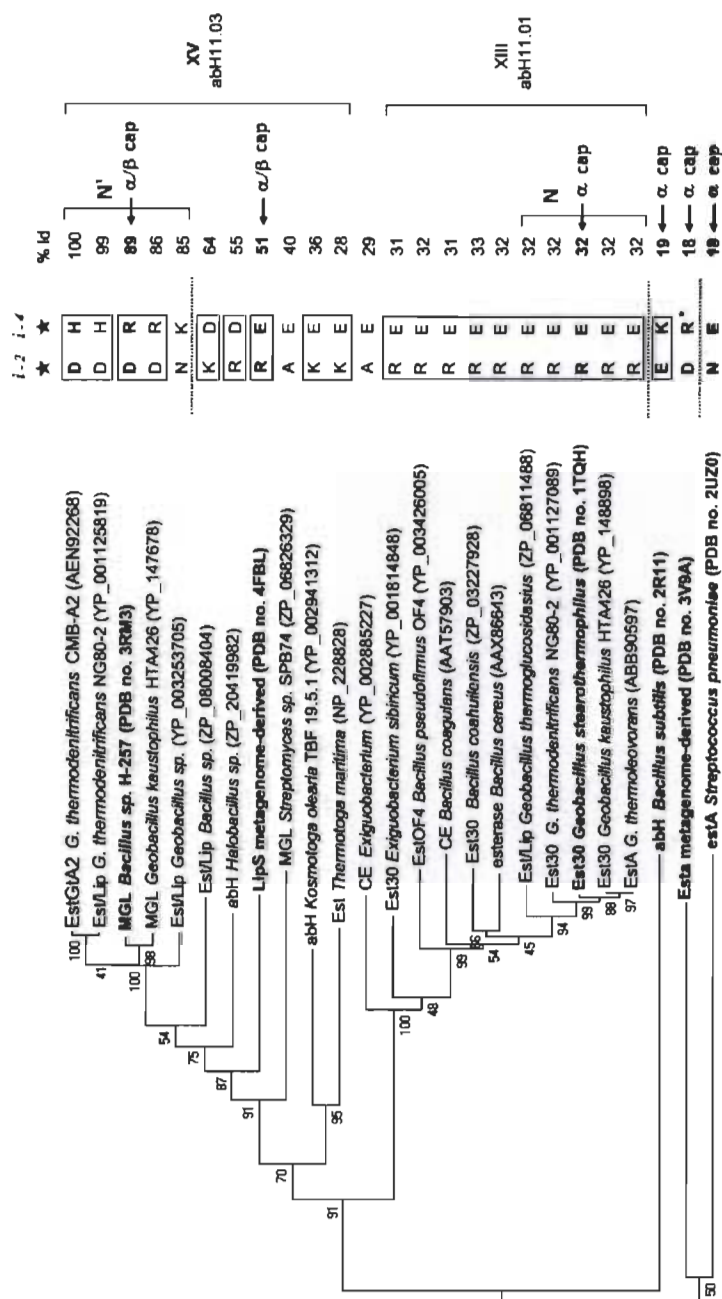


Figure 1.12 Alignement de séquences multiples montrant l'évolution d'un pont salin près du site actif d'EstGtA2.

Des enzymes provenant de plusieurs familles (XV, XIII) y sont présentés ainsi que les divisions entre les sous-familles N et N'. Les colonnes présentent les acides aminés impliqués dans ce pont (pour EstGtA2; la première ligne; D et H) ainsi que le pourcentage d'identité par rapport à EstGA2. Les lignes pointillées montrent les inversions de polarité (voir la figure 1.1 pour les résidus chargés négativement et ceux positivement). L'échelle représente la distance en phylogénie entre chaque enzyme. Source : Charbonneau, 2013.

Ce genre d'information est détectable rapidement lors de l'étude de la séquence de l'ADN d'un génome, et permet de supposer que le gène en question encode une estérase proche de celles ayant la même signature. Une autre façon d'aider à la compréhension du lien qu'il existe entre la séquence en acides aminés et le comportement enzymatique est l'utilisation de la mutagenèse dirigée, une méthode dans laquelle on va générer des mutants dont le choix est basé soit sur un modèle ou sur la structure 3D. Malheureusement, la génération de mutants basés sur cette méthode comporte des risques comme la perte d'activité ou de la stabilité. Dépendamment des informations détectées, on arrivera à prédire les propriétés d'enzymes encodées, sans les produire ou les caractériser.

Donc, dans le futur, l'annotation génomique nous permettra de prédire, à partir de la séquence d'ADN des régions codantes, le type de repliement, les propriétés enzymatiques et les applications potentielles des protéines encodées dans un génome. Le défi actuel est la génération d'information clé. Notre équipe focalise sur l'élucidation de signatures dans les estérases procaryotes et s'inscrit dans un effort à l'échelle planétaire impliquant des centaines de laboratoires étudiant autant de familles de protéines. Le sujet de ce mémoire s'inscrit dans notre démarche, en étudiant le comportement d'une nouvelle estérase (EstGtA2) et en tentant d'élucider des relations séquence-structure-propriétés dans cette enzyme.

1.9 EstGtA2

EstGtA2 est une enzyme qui a été caractérisée pour la première fois dans notre laboratoire par mon collègue David Charbonneau (Charbonneau, Meddeb-Mouelhi et Beauregard, 2010; Charbonneau et *al.*, 2012a), qui avait comme premier mandat de caractériser de nouvelles enzymes provenant de souches thermophiles. *Geobacillus thermodenitrificans* est un organisme isolé à partir de compost et qui croît à des températures variant entre 50 et 60 °C. Ces températures impliquent donc que ces protéines sont stables à ces températures. Lors de l'identification de la souche dans Genbank (GQ293454), un gène a particulièrement attiré l'attention de mon collègue. Ce gène (YP_001125819), qui codait normalement pour une estérase (selon le type de

séquence), n'avait jamais été cloné, et l'enzyme jamais caractérisée. Brièvement, le gène fut cloné dans un système procaryote (*Escherichia coli* M15) à l'aide d'un vecteur recombinant et la protéine surexprimée dans le but d'être étudiée.

EstGtA2 est une enzyme de 249 acides aminés qui doit son nom directement à sa fonction (**Estérase**), à l'organisme dans lequel elle a été découverte (*Geobacillus thermodenitrificans*) ainsi qu'au nom qui avait été donné à cette souche en laboratoire (CMB-A2). Cette enzyme permet à la bactérie de dégrader plusieurs types de substrats lipidiques. La température et le pH optimaux de l'enzyme ont été déterminés (50 °C, pH 8) et un modèle par homologie été généré à partir d'une protéine homologue (Est30, 30 % d'identité, PDB 1TQH). Comme le but de nos recherches est d'incorporer les enzymes dans les milieux industriels, mon collègue s'est dès lors attardé à comprendre, ce qui confère une stabilité dans la structure d'EstGtA2. Ces recherches ont, entre autres, étudié l'effet du pH sur l'activité et la stabilité de l'enzyme, en plus d'étudier le rôle de plusieurs ponts salins présents dans l'enzyme (Charbonneau, Meddeb-Mouelhi et Beauregard, 2010; Charbonneau, Meddeb-Mouelhi et Beauregard, 2012b).

1.10 Objectif

L'objectif général de ce travail de maîtrise était de déterminer et de comprendre l'impact d'une variation de pH en plus d'une variation de force ionique sur la thermostabilité de la carboxylestérase EstGtA2 provenant de *Geobacillus thermodenitrificans*. De plus, des essais de cristallisation d'EstGtA2 ont été faits en parallèle, dans le but d'élucider la structure tridimensionnelle de l'enzyme par diffraction de rayon X en collaboration avec le Dr Albert Berghuis de l'Université McGill.

CHAPITRE II

LE RÔLE DES RÉSIDUES CHARGÉS EN FONCTION DU PH DANS LA DÉNATURATION THERMIQUE DE LA CARBOXYLESTÉRASE ESTGTA2

Le contenu de ce chapitre est présenté sous forme de manuscrit qui a fait l'objet d'une soumission dans le journal scientifique « Journal of Biochemistry ».

2.1 Contribution des auteurs

L'ensemble des expériences à la base de cet article et la majorité des figures ont été réalisées par Jessica Moisan, qui a aussi écrit la première ébauche du manuscrit. Les analyses et figures concernant la structure tridimensionnelle et les potentiels électrostatiques ont été réalisées par David M. Charbonneau. Ces recherches et l'ensemble des expériences ont été menés en étroite collaboration avec Fatma Meddeb, David Charbonneau et Marc Beauregard qui ont participé à la discussion des résultats et à l'écriture finale du manuscrit.

2.2 Résumé de l'article

Les carboxylestérases sont des enzymes ayant une grande applicabilité dans les procédés industriels. Les carboxylestérases provenant de souches thermophiles possèdent d'autant plus d'intérêt puisqu'elles résistent naturellement à des conditions de températures extrêmes. Cependant, la plupart de ces enzymes qui sont bien caractérisées doivent être mutées génétiquement afin de répondre aux conditions spécifiques des procédés dans lesquels elles sont utilisées (force ionique, pH). Comme la séquence en acides aminés a un impact direct sur le rôle et la stabilité d'une enzyme, la compréhension de cette relation devient primordiale. Afin de cibler les facteurs clés impliqués dans la stabilité d'EstGtA2, la dénaturation de l'enzyme dans différentes conditions de pH et de

forces ioniques a été caractérisée. Des tests de dichroïsme circulaire, de diffusion dynamique de la lumière et de thermostabilité ont été réalisés et les résultats ont mis en évidence les différents types d'agrégation et d'association subis par EstGtA2 pendant sa dénaturation. Plus précisément, un état non-natif où les structures secondaires sont conservées au-delà de 100 °C a été observé près du pH neutre (pH 7-9), dépendamment des concentrations ioniques tandis que d'autres états de dénaturation ont été identifiés à ces extrêmes. Les chemins menant aux différents états de dénaturation en pH acide et basique sont directement contrôlés par le pH tandis que la force ionique permet de moduler l'association autour du pH neutre, démontrant ainsi le rôle des interactions ioniques en surface d'EstGtA2.

2.3 Article scientifique

Charged residues control denaturation and protein association pathways in carboxylesterase EstGtA2

Jessica Kelly Moisan^{1,2}, Fatma Meddeb-Mouelhi^{1,2}, David M. Charbonneau^{2,3} and Marc Beauregard^{1,2*}

¹ Centre de recherche sur les matériaux lignocellulosiques, Université du Québec à Trois-Rivières, Trois-Rivières (Québec) Canada, G9A 5H7

² PROTEO (Quebec network for research on protein function, engineering and applications), Université Laval, Québec (Québec) Canada, G1V 0A6

³ Département de chimie, Université de Montréal, Montréal (Québec) Canada, H3T 1J4

Running title: Denaturation and aggregation of EstGtA2

*** Corresponding author:**

Marc Beauregard, Ph.D.

CRML, Université du Québec à Trois-Rivières, C.P. 500, Trois-Rivières (Québec) Canada, G9A 5H7

Tel: (819) 376-5011 # 3354

Fax: (819) 376-5084

E-mail: marc.beauregard@uqtr.ca

Summary

Understanding the relationship between enzymatic stability and the amino acid sequence encoding such enzymes is of utmost importance. Here we thoroughly characterized the behavior of the carboxylesterase EstGtA2 from *Geobacillus thermodenitrificans* during thermal denaturation at different pH with various salt concentrations. EstGtA2 was characterized by circular dichroism regarding conformation and thermal stability, by dynamic light scattering for detection of association/aggregation, by enzymatic assays for activity, and by monitoring recovery of activity after heat treatments. Our investigation revealed a particular dependence between aggregation/association and preservation of secondary structures upon heating in EstGtA2. At pH 7 and 8, depending on salt concentration, a folded but non-native associated state characterised by an apparent particle size of 300 nm resisted secondary structure unfolding up to 95°C. The paths leading to various aggregative states were found to be controlled by pH (depending on proximity to pI) and to a lesser extent, ionic strength, suggesting that ionic interactions at the surface of the protein are responsible for behavior of EstGtA2. The various paths available to EstGtA2 could be important for protection of *Geobacillus thermodenitrificans* when exposed to heat stress. The understanding and/or control of these paths would allow for optimal use of EstGtA2 in industrial processes.

Keywords: Thermal denaturation, circular dichroism, dynamic light scattering, protein aggregation, protein association.

Introduction

Lipolytic enzymes (esterases and lipases) are ubiquitous in nature, being found in plants, animals and microorganisms. They belong to the serine hydrolase family, display a common α/β hydrolase fold and share a similar active site consisting of a highly conserved catalytic triad (Ser-Asp/Glu-His) [1,2]. Lipases (EC 3.1.1.3) are defined as enzymes that catalyse hydrolysis of acylglycerols with long chains (≥ 10 carbon atoms) while closely related carboxylesterases (E.C 3.1.1.1) are defined as enzymes that catalyze the hydrolysis and the synthesis of short chains (aliphatic or aromatic ester smaller than 10 carbon atoms) [3]. Lipolytic enzymes from prokaryotes have been originally grouped into eight families by Arpigny and Jaeger according to conserved amino acid sequence motifs and biological properties [4]. The ESTHER database rallies lipolytic enzyme families extending from the original Arpigny and Jaeger's classification (AJ) and from the Lipase Engineering Database (LED) [5, 6]. LED sorts enzymes according to their sequence, structure and function, resulting in 38 homologous families and 16 super families [6].

Nowadays, many microbial esterases are used in industrial applications. Unfortunately, many enzymes fail to perform under harsh conditions and are not suitable to fulfilling specific industrial needs [7, 8, 9]. Like most proteins, enzymes are sensitive to reaction conditions and are prone to denaturation and aggregation under certain conditions. Thermal denaturation inactivates enzyme but protein aggregation may have more important consequences. In the industrial field, prolonged exposure to heat may induce amorphous aggregation, which prevents refolding and recovery of activity regardless of temperature decrease. Aggregated enzymes have to be replaced, decreasing process profitability. Protein aggregation is a major issue for human health too: it is involved in degeneracy illnesses such as Alzheimer and its occurrence in therapeutic protein formulation may have dramatic impacts [10-12]. Therefore, understanding the mechanisms underlying protein aggregation is an important topic for industrial, environmental and medical purposes.

Our current understanding of protein aggregation is far from satisfactory, but considerable progress has been achieved over the past decades. Recent studies have shown that aggregation is mainly dependent on 3 factors: experimental conditions, stability of the secondary and tertiary structures, and the amino acids sequence itself [10]. Stressful environmental conditions can lead to the exposition of specific amino acid sequences labeled as “aggregation-prone sequences” which are normally hydrophobic and/or weakly charged. During thermal denaturation, the unfolding of the protein chain leads to the exposition of hydrophobic residues to solvent (initially buried in the hydrophobic core) often promoting irreversible aggregation. A recent study demonstrated that charged amino acid are also to be considered in such hot spot sequences because of their essential conformational stability role [13]. Thermally induced protein aggregation depends on the protein sequence itself, on protein concentration (solubility being dependant on concentration) and on the environmental parameters prevailing when it unfolds. In particular conditions, the unfolding of a protein chain may cause the subunits to aggregate in a totally disordered manner (amorphous aggregates) or to associate in an highly organized fashion, such as the β -amyloid fibers associated to Alzheimer’s disease [12]). Protein denaturation may also lead to reversible subunit unfolding and association, characterised by a soluble unfolded state and/or by soluble particles, in which every unfolded subunit can refold when the conditions allow it [14]. Those reversibly denatured enzymes can recover from moderate and short term temperature variations which are frequent in industrial processes and therefore are more profitable biocatalysts [13]. Pathways by which enzymes unfold and aggregate can be investigated by characterising new proteins of variable thermostability [15] or by comparison to related but modified proteins (mutagenesis, directed evolution) [16].

Previously we reported on EstGtA2, an esterase associated with AJ family XIII.2 (and LED family abH 11.03). We have investigated pH-dependent heat denaturation of EstGtA2 and found that its reversibility (after thermal denaturation) was only possible under neutral and alkaline pH. By contrast, acidic conditions lead to irreversible thermal denaturation, possibly due to aggregation [17]. This observation triggered the identification of key charged residues involved in salt bridges which would be important

for the thermostability of EstGtA2 and closely related enzymes of family XIII.2 (or LED abH11.03) [17,18]. EstGtA2 was shown to harbour five exclusively conserved salt bridges. When the bridges were disrupted via mutagenesis, dramatic impacts were reported, such as an important decrease in melting temperature (down from its native T_m of approximately 70°C) or total inactivation of the enzyme when accumulating more than four residue mutations [17]. EstGtA2 ability to recover from heat stress was also investigated and pH was found to modulate thermostability, as recovery from heat stress was allowed at alkaline pH but not under acidic pH. As expected from previous literature [19-21], surface ionisable side chains in EstGtA2 do play a role in determining thermostability. However, the impact of surface charges on unfolding (as revealed by T_m) on one hand, and on aggregation (due to prolonged exposure to heat during thermal treatments) on the other hand, were not delineated for EstGtA2. Clearly, understanding how these phenomena control aggregation is vital for developing a strategy aiming at improving the industrial applicability of EstGtA2 and will help do the same for related enzymes.

The importance of expanding the knowledge of lipases is not limited to industrial or medical applications. Rapid and high Throughput DNA sequencing being easily accessible, the ability to make sense of the ever-increasing genetic information (genomic annotation) is now mainly dependent on the generation of biochemical knowledge developed by enzymologists with every new enzyme being characterised. Structural and biochemical characterisation of such new proteins allowed for identification of key structural features that are believed to be responsible for thermostability. Presence of additional salt bridges, increased compactness in the hydrophobic core and the formation of surface charge networks by ionisable residues were shown to be hallmarks of thermophilic enzymes [19, 20, 21]. The elucidation of aggregation prone signatures and their inclusion in genomic annotation lexicon has not received much attention despite its importance.

In this paper, we studied how EstGtA2 respond to pH and ionic strength variations. These two experimental parameters were selected because of their possible impact on

EstGtA2 surface charges and salt bridges, and also because of their importance in determining enzyme performance in an industrial setting. To this end, we carried out biophysical and biochemical characterisations of carboxylesterase EstGtA2 and identified parameters that control denaturation and aggregation pathways of this enzyme. It appears that pH controls aggregation pathways for EstGtA2 and that salt content modulates the dominant impact of pH.

Materials and methods

Strains, plasmids and reagents

The expression system consisted of *E. coli* BL21 Gold(DE3)pLysS in which the pET28-EstGtA2 plasmid construct was cloned for over expression of EstGtA2 (8). The buffers used were 20 mM sodium phosphate (pH 5-9) and 20 mM sodium carbonate (pH 10-12). The ionic strength was adjusted to 50 mM or 120 mM (final sodium salt concentration) by the addition of NaCl (30 mM or 100 mM) in the respective buffer.

Protein expression and purification

Growth medium (850 mL TB) containing kanamycin (50 µg/mL) was inoculated with 850 mL of saturated overnight culture at 37°C. The cells were grown at 37°C with shaking at 200 rpm until the culture reached an optical density (OD_{600 nm}) of 0.8. The protein expression was then induced by the addition of 0.5 mM IPTG (final concentration) followed by an 8 h incubation at 25°C under agitation (200 rpm). Cells were harvested by centrifugation at 4000 rpm for 15 min (4°C) and stored at -80°C. Pellets were then thawed, resuspended in 15 mL of lysis buffer (50 mM sodium phosphate pH 8, 2 mM imidazole, 300 mM NaCl, 0.1 mM PMSF) and the cells lysed by sonication (6 cycles x 60 s at 60 pulses/sec). The lysate was clarified by centrifugation at 15,000 rpm for 30 min (4°C). The clarified cell lysates containing His-tagged recombinant EstGtA2 protein was applied onto Ni-NTA resin beads for 1h at 4°C under rotated agitation. The nickel resin was washed with 60 mL of lysis buffer and the protein was recovered

with 10 mL of elution buffer (50 mM sodium phosphate pH 8, 250 mM imidazole, 300 mM NaCl). The purified EstGtA2 was dialyzed O/N at 4°C against 4 L of sodium phosphate buffer (20 mM, pH 8) using 3.500 kDa MWCO dialysis membranes. Protein concentration and production yield (80 mg/L of culture) were determined via Bradford method using bovine serum albumin (BSA) as standard. Protein purity was verified by SDS-PAGE.

CD spectroscopy

Thermal unfolding and CD spectra of EstGtA2 were recorded with a Jasco J-720 spectropolarimeter equipped with a Peltier temperature controller under nitrogen atmosphere using protein concentration of 0.5 mg/mL under different pH and ionic strength conditions (20, 50 and 120 mM sodium salt). For thermal denaturation in the far UV (190-240 nm) and near UV (250-310 nm), transitions were monitored by the variation of CD signal at 222 and 272 nm respectively, using a 0.01 and a 1 cm quartz cell respectively. Temperature was ramped from 25 to 100°C using a Peltier temperature controller at a rate of 0.83°C/min. Denaturation transitions were analysed using a nonlinear least squares fit assuming a two-state model ($F \leftrightarrow U$) characterised by pre and post-transitional regions. Extrapolation and least-square analysis from the pre- and post-transition regions allowed determination of the equations for y_F and y_U . The equilibrium constant was determined at each temperature and the thermodynamic parameters were determined according to the following equations [22]:

$$\text{Eq 1. } f_U = (y_F - y) / (y_F - y_U)$$

$$\text{Eq 2. } K = f_U / (1 - f_U) = (f_U / f_F) = (y_F - y) / (y - y_U)$$

The free energy change associated with thermal unfolding was calculated at each temperature by the following equation:

$$\text{Eq 3. } \Delta G = -RT \ln(K) = -RT \ln(y_F - y) / (y - y_U)$$

The apparent T_m values were then determined by plotting ΔG as a function of temperature where $T_m = T$ at $\Delta G = 0$, the midpoint of thermal unfolding curves, where $\Delta G(T_m) = 0 = \Delta H_m - T_m \Delta S_m$. Spectra were corrected for buffer signal and conversion to Mol CD ($\Delta \epsilon$) was performed with the Jasco Standard Analysis software. The appearance of protein solution after CD thermal denaturation was determined by a naked-eye observations through the quartz cell.

Dynamic light scattering

Dynamic light scattering (DLS) experiments were carried out using a Dynapro Nanostar (Wyatt Technology) and protein aggregation/association was monitored during denaturation. Protein concentration was adjusted to 0.5 mg/mL under different pH values and ionic strength conditions. Protein solutions (10 μ l) were first centrifuged to remove any dust particles and then transferred into a clean and dust-free micro quartz cell (1 μ l-DPN from Wyatt technology). Protein samples were covered with one drop mineral oil to avoid evaporation during heat. Denaturation was induced at a heating rate of 0.8°C/min and measurement were set to record as a mean of 3 acquisitions of 15 seconds at every degree. All samples were measured in triplicates and the mean value was reported.

Isoelectric focusing

Isoelectric focusing was performed using the Mini IEF Cell Model 111 (Biorad 170-2975). For the determination of EstGtA2 isoelectric point (pI), 5 μ g of protein were run on a 0.4 mm thick polyacrylamide gel prepared as suggested in the Mini IEF instruction manual (170-2976) with minor modifications. The gel was prepared with adjusted concentration of cross linker (T=5% and C=4%), where “T” represents the weight percentage of total acrylamide and “C” the proportion of bis-acrylamide as percentage of total monomer. Ammonium persulfate concentration was increased to 0.7 mg/mL and solution degassing time reduced to 4 minutes. Ammonium persulfate, riboflavin-5'-phosphate and TEMED were used for polymerisation catalysis. The ampholyte solution Bio-lyte 3/10 (Biorad 163-1112) was used to establish a pH gradient (3-10) and a

fluorescent lamp was used to accelerate the polymerization. Migration of EstGtA2 was monitored with visible marker proteins of known pI (IEF standard, Biorad, # 161-0310). Focusing was carried out at 100 V for 15 minutes, 200 V for 15 minutes and finally 450 V for 30 minutes or until migration had ceased. Protein bands were detected by fixing, staining and destaining as recommended by the supplier.

Enzyme activity measurements

Carboxylesterase activity was assessed spectrophotometrically by measuring the amount of *p*-nitrophenol (*p*NP) released after hydrolysis of *p*NP-caprylate (*p*NP-C8) as substrate. The substrate was prepared in 100% acetonitrile solvent and EstGtA2 was diluted in phosphate buffer. The reaction mixture had a final volume of 200 μ L with final salt concentrations of 20 to 120 mM. Each reaction contained *p*NP-C8 (100 μ M), Tween-20 (0.04%) and 1 μ g of the purified enzyme. Enzymatic reactions were performed at 25 and 65°C in 96-well microplates. For all enzymatic activity measurements, the substrate and the enzyme were preincubated during 15 minutes at the appropriate temperature before starting the enzymatic assay. The production of *p*NP was monitored at 405 nm using a SynergyMx microplate reader (BioTek). The molar extinction coefficient ($M^{-1} cm^{-1}$) of the product (*p*NP) was determined for each set of conditions to quantify the amount of *p*NP (μ mol) released during the enzymatic reactions. The background hydrolysis of the substrate was subtracted by using a control without the addition of the enzyme. The amounts of *p*NP released (μ mol) as a function of time (min) were used to define the initial velocities (μ mol/min) of enzymatic reactions. The activity is then reported in U/mg where one unit is defined as the rate of *p*NP-caprylate hydrolyzed (μ mol/min) per milligram of pure enzyme.

Activity recovery after heat stress

The residual activity of EstGtA2 after heating and cooling was quantified (using the same method as described in the above section) after boiling the enzyme for 5 and 15 minutes. The enzyme (0.01 mg/mL) was prepared in phosphate buffer solution with

varying pH condition (pH 6, 7 and 8) and with different final ionic strength (20, 50 and 120 mM sodium salt). The enzyme was cooled down to 25°C for 20 minutes before measuring its residual activity. The residual activity was compared to the non-heated control made under the same conditions (pH and ionic strength). The reaction (200 μ l) contained 20 μ l of *p*NP-caprylate (100 μ M), 80 μ l of phosphate buffer supplemented with 0.04% Tween 20 and 1 μ g of EstGtA2 (100 μ l of cooled down enzyme).

Model of EstGtA2 and mapping of surface electrostatic potential

The model of EstGtA2 was created by homology modeling using (PDB 3RM3) as template (89% residue identity) as previously reported [17]. Continuous surface electrostatic potential was computed on EstGtA2 model using the Adaptive Poisson-Boltzmann Solver (APBS) software package which solved the Poisson-Boltzmann equation describing the electrostatic interactions between solutes in salty aqueous media [23,24]. The electrostatic potential was solved for EstGtA2 (dielectric constant = 2.0) in water (dielectric constant = 78.0) at pH 5, 7 and 8 at 80°C, without salt (0 mM monovalent counterions) or in the presence of 120 mM of monovalent counterions (equivalent NaCl).

Results

Conformational and thermodynamic analyses of secondary structures

The thermal denaturation of EstGtA2 was investigated at various pH and ionic strength in order to understand its pH-dependent thermostability profile. To this end, the concentration of sodium salt was modulated by adding NaCl to the different buffers used (20 mM sodium phosphate or 20 mM sodium carbonate depending on target final pH) leading to final sodium salt concentrations of 20, 50 and 120 mM. As shown in Figure 1, melting temperatures that were computed from complete transitions were maximal at pH 7. Results suggest that pH and not ionic strength was critical for secondary structure denaturation. Ionic strength did have an impact however on the completeness of

denaturation. At pH 7 and under low ionic strength (20 mM), EstGtA2 secondary structures remained folded up to 95°C, making it impossible to calculate an apparent T_m . At pH 7 and 20 mM salt, the melting temperature was estimated to be at least 30°C higher than calculated for pH 6 and 8 at similar ionic strength, where apparent T_m values of 74.1 and 68.5°C were respectively calculated. When ionic strength was increased to 50 and 120 mM melting temperatures of EstGtA2 were calculated and compared to the results mentioned above (Fig. 1). At pH 7, the increased ionic strength resulted in the complete denaturation of EstGtA2 secondary structures and melting temperatures of 77.2 and 74.1°C were calculated (for 50 and 120 mM salt, respectively). As pH moves from 7 to 9, salt concentration had to be increased in order to promote resistance of secondary structure to thermal unfolding. At pH 8, this was observed at 50 and 120 mM of salt while at pH 9, absence of transition was detected using a 120 mM total salt concentration. This apparent resistance of secondary structures to high temperature demands further investigation.

Figure 1

Table 1

Enzyme solutions were also visually inspected after thermal runs in order to detect aggregation. At low ionic strength (20 mM), thermal denaturation led to appearance of large and visible aggregates for pH 5 and 6, while it led to clear solutions for pH values ranging from 8 to 11 (see Supplementary Materials, Figure S1). At pH 7 (20 mM salt), an intermediate level of turbidity was observed (marked as “cloudy” in Table 1), clearly different from what was observed at other pHs with similar salt concentration. The same cloudy appearance was observed at increased ionic strength conditions (120 mM) for pH 8 and 9. This cloudy appearance was observed under conditions that prevented loss of CD signal (suggesting resistance to secondary structure unfolding).

Spectra and melting curves associated with these different aggregative states shed some light on their significance. Figure 2 shows typical far UV CD spectra associated with low pH induced aggregation, high pH induced solubility, and inability to induce

secondary structures denaturation at pH 7, 8 or 9 (depending on ionic strength). Boiling of EstGtA2 treated at pH 5 lead to nearly zero CD signal, which is expected from turbid, aggregated solutions. Then at pH 8 (20 mM salt, see red curves) high temperature lead to a typical soluble denatured state (minima shifting from 215 nm to 200 nm) [25]. Under conditions where no transitions were recorded (pH 7, 20 mM salt, Table 1) the signal remains unchanged (black dashed line vs black solid line). Melting curves associated with such results are available in Supplementary Materials Figure S1.

Figure 2

Conformational and thermodynamic analysis of tertiary structures

The thermal unfolding of tertiary structures was monitored by following the near-UV signal at 272 nm in order to estimate the melting temperature of the hydrophobic core (referred to T_m^3) at different pHs (6, 7 and 8) and salt concentrations (20 and 120 mM), as shown in Table 1. Results obtained at pH 7 (20 mM) indicated that EstGtA2 lost the compactness of its hydrophobic core at 72.3°C (see Fig. S2), at a temperature where the secondary structures appeared folded (Fig. 2).

Table 2

For all pH tested, this near UV CD investigation suggests that EstGtA2's core stability is not influenced, to a large extent, by pH or salt, with values ranging from 66 to 72°C for all conditions tested here (pH 6 to 9, 20 and 120 mM salt, see Table 1), including 120 mM salt. This tertiary structure denaturation investigation suggests that a conformational change occurred for all pH and all salt concentrations tested when EstGtA2 was heated to approximately 70°C, regardless of secondary structures unfolding (or not, depending on conditions). If we had to estimate the “expected” secondary structures T_m for pH 7 at low salt or for pH 8 at 120 mM salt (bars ending with stars in Figure 1), such values would be in the 70-75°C range. Slightly higher but close to T_m^3 values reported in Table 1 under same conditions. This comparison strongly suggests that

the phenomenon leading to an unexpectedly high stabilisation of secondary structures (T_m of 95°C or above, as indicated by stars in figure 1) did not lead to any concomitant stabilisation of hydrophobic core in EstGtA2 (which melted in the 66 to 72°C range) and is not driven by destabilisation of EstGtA2 protein core either.

Analysis of protein aggregation by dynamic light scattering

Dynamic light scattering was used for the detection of association or aggregation promoted by exposure to high temperature while performing thermal denaturation. The hydrodynamic radius (R_h) of particles in the protein samples was monitored for pH 6, 7 and 8 at two salt concentrations (20 and 120 mM). Dynamic light scattering investigation suggested the formation of three aggregated (or associated) states, as indicated by particle R_h (Table 1). The most obvious correlations were observed between pH and aggregates size. Large aggregates dominated by 2000 nm component were detected under acidic condition (pH 6 and below) while small particles (8 nm) were detected after thermal denaturation under high pH condition (pH 8 and above). Under pH 7 (intermediate pH value), a third type of aggregate was detected ($R_h = 300$ nm) depending on salt concentration in addition to the 8 and 2000 nm particles mentioned above. At low salt concentration (20 mM), a nearly equal mix of small and medium size aggregates was observed at pH 7 (8 nm and 300 nm). With 120 mM salt, larger aggregates were observed (2000 nm particles accounted for 90% of mass content). At pH 8 (20 mM salt) mainly small aggregates were observed (90% of 8 nm species based on mass). CD measurements supported the existence of a soluble denatured state at pH 8, 20 mM salt while it revealed amorphous aggregation at pH 5 (loss of signal at high temperature due to visible aggregation (Fig. 2)). A mixture of large and medium aggregates was observed at pH 8 under higher ionic strength conditions (120 mM), where larger aggregates (300 and 2000 nm) were observed. Medium and small aggregates were detected at pH 9 at 120 mM salt. A simple visual inspection corroborated results from DLS analyses for all conditions. Observation of protein samples after CD denaturation experiments confirmed that acidic conditions induced visible aggregation and high pH preserved solution transparency (Supplementary Materials Fig. S3), supporting our interpretation of DLS experiments.

Clearly, CD and DLS results suggest that mainly pH determine the aggregation or association states in EstGtA2 after thermal unfolding, but that under moderate pH values (7 to 9), ionic strength modulates pH impact on EstGtA2 behavior.

The concomitant observation of a mid-size aggregated content (300 nm species leading to visible cloudiness) and protection of secondary structures from heat denaturation points to a relationship between both phenomena. We addressed this hypothetical relationship between aggregation and preservation of far UV CD signal by repeating the thermal unfolding of secondary structures in the presence of urea, a well-known aggregation inhibitor. Melting temperatures measured in the presence of 2 M urea were compared to their control values (measured without urea, Table 2). The addition of urea allowed for complete denaturation of secondary structures in EstGtA2 at pH 7 (20 mM salt), with a T_m of 66.7°C (compared to an apparent T_m higher than 95°C in absence of urea). This important drop in T_m cannot be ascribed to destabilisation of protein conformation by urea since moderate impacts was observed. The impact of 2 M urea on melting temperature measured under pH 6 and 8 lead to a decrease in T_m by about 6°C (Table 2). Therefore, the impact of urea on EstGtA2 behavior at pH 7 suggests that protein association or aggregation plays a role in the apparent thermal resistance of secondary structures observed here under specific conditions.

Determination of isoelectric point

Considering that salt and pH have an obvious impact on EstGtA2 behavior, the determination of the isoelectric point (pI) provided valuable information on what promoted aggregation. Using isoelectric focusing, a pI of 4.9 was measured for EstGtA2 (Fig. S4). This explained the behavior of EstGtA2 at pH 5 or 6, where large aggregates were generated by heating at all salt concentrations investigated. Under higher pH overall repulsion between proteins with a pI of 4.9 should prevail, avoiding aggregation.

Figure 3

Effect of pH and salt concentration on EstGtA2 activity

Enzymatic activity measurements were performed to further understand the effect of pH and ionic strength on the protein behavior at two temperatures (25 and 65°C, the upper temperature limit available when performing assays in a Biotek plate reader). *p*NP-C8 was used as a substrate since EstGtA2 showed highest activity using this substrate in an earlier study [20]. At 25°C, increasing salt concentration from 20 to 120 mM resulted in an increased activity for all tested pHs (6, 7, and 8). The same trend was observed when enzymatic activity assays were performed at 65°C (Fig. 3). These results indicate that EstGtA2 is more active in the presence of 120 mM salt and at 65°C which is slightly below denaturation temperatures reported in Figure 1 and Table 1. No correlations can be inferred between activity dependence on pH or salt concentration and similar dependence on denaturation temperature or formation of various aggregated states during heating observed above. This suggests that enzyme denaturation, aggregation or association did not proceed at 65°C or that exposure to such temperature was too short to induce activity loss (compared to the duration of full thermal ramps used for CD and DLS experiments).

Figure 4

Investigation of activity recovery after heat treatment

The ability to recover from a heat stress in enzymes is a function of denaturation temperature, aggregation and ability to refold. The impact of pH and ionic strength on recovery of activity after heat stress (sometimes referred to as thermostability) was investigated for EstGtA2. To this end, the enzyme was exposed to pH 5 to 8 at 20, 50 and 120 mM salt concentrations and heated for 5 and 15 minutes at 100°C. Activity measurements were carried out after cooling down at 25°C (Fig. 4). At pH 5 and 6, there was no residual activity after heat stress, regardless of salt concentration used (data not shown). It is suggested that aggregation rather than thermal stability of EstGtA2 leads to absence of activity recovery. The best residual activity was measured at pH 7 and pH 8 after 5 minutes of boiling, while it decreased dramatically when boiling was extended to 15 minutes. The conditions that induced the highest denaturation temperatures and limited

aggregation (when available from Fig. 1 and Table 1) were the ones that promoted maximal recovery. Here the apparent resistance of secondary structures toward heat denaturation under specific conditions (*e.g.* pH 7, 20 mM salt, pH 8 with 50 mM salt) did not seem to have any noticeable impact on activity recovery. The limited heating time used for recovery experiments (compared to denaturation heating ramps used for CD and DLS (1 hour)) may not be sufficient to populate the 300 nm species apparently associated with secondary structures resistance to denaturation.

Discussion

Previous work by Charbonneau et al. (2013) pointed to the role of exposed ionisable residues in determining thermostability in EstGtA2 [17]. Here we focused on the impact of pH and salt concentration on a variety of conformational parameters, as described in Tables 1 and 2. Overall, this investigation reveals the dominant impact of pH on conformational stability, on the size of aggregates after prolonged heating and on the ability to recover from heat stress (thermostability). Maximal activity recovery after heating was observed at pH 7 and 8, where maximal T_m values were also recorded (Tables 1 and 2) and where aggregates of intermediate size were detected (via DLS and suggested by simple visual inspection). When pH increased to 9 or above, the aggregate decreased in size, but T_m did decrease too. When EstGtA2 was incubated at pH 5 and 6, T_m values were close to maximal T_m (pH 7, 50 mM salt), but large aggregates were detected. Under these conditions, the enzyme did not recover from a heat stress followed by cooling. We suggest that under acidic pH, recovery from heat stress is prevented by irreversible and amorphous aggregation, while under alkaline pH, it is limited by lack of conformational stability as shown by small T_m values (45 to 65°C).

Under acidic conditions (pH 5 and 6), where pH is near the isoelectric point of EstGtA2 (pI of 4.9 measured here for the first time), heated EstGtA2 denatured and aggregated as large particles (2000 nm). At pH 5 and 6, EstGtA2 subunits carry little net electrical charges and do not repulse each other as much as they would at higher pH. When an environmental stress, such as heating, is applied to the enzyme, the torsion

induced in the tertiary structure and the proximity between subunits promote buried amino acids to get exposed. It induces those newly exposed amino or hydrophobic patches of many subunits to command irreversible/amorphous aggregation. Accordingly, salt addition had no effect for all properties measured under acidic pH except for a limited impact on T_m (which varied by 0.3 °C or less depending on salt concentration). Clearly, the addition of ions in solution did not prevent hot spots triggered aggregation. Thermostability measurements performed at pH 6 support this hypothesis, as no residual activity was detected for the three tested ionic strengths. Lack of recovery from heat stress cannot be ascribed to a large decrease in T_m here (in the 70°C range), so it is suggested that large aggregates formed under acidic pH are amorphous and prohibit refolding. The path leading to such aggregates under acidic pH is discussed below.

Under alkaline conditions, heated EstGtA2 formed soluble particles (8 nm), which were apt for refolding. Under such conditions, the surface of the protein is expected to be negatively charged overall, promoting the repulsion between EstGtA2 subunits and preventing the aggregation into large aggregates. Increased pH led to an important decrease in conformational stability. One can hypothesize that accumulation of negative charges may destabilise each individual enzyme, or that a particular ionic status at key positions is necessary for optimal stability. Increasing salt did little to stabilise secondary structures, giving support for the second hypothesis.

Our study revealed a pH dependent turning point between the two patterns of denaturation discussed above. This turning point, in which the enzyme retains its secondary structures and where many subunits associated into a folded but non-native state, is modulated by salt. Secondary structures were stabilized when 300 nm particles appeared, leading to a particular cloudiness in solution, under conditions where tertiary structure was compromised. At pH 7 this non-native but folded associative intermediate was populated in the presence of low salt. When pH was shifted further away from pI, then more salt had to be increased in order to access this state. This clearly indicates that EstGtA2 aggregation path during thermal denaturation is directly modulated by the shielding of surface charged amino acids by an increased ionic strength. Overall charge is

determined by pH, but shielding clearly dictated the most adequate denaturation-aggregation path.

Salt was increased from 20 to 120 mM in order to verify the impact of charge neutralisation or masking on stability and aggregation. It appears that at the concentration used, salt content had a minor impact for most parameters studied here. It did not lead to any clear modification of T_m (as revealed by both near and far UV CD) and recovery from heat stress. The impact of salt concentration was not visible under pH 5, 6, 10 and 11, where pH clearly controlled T_m and aggregate particle size. Salt appeared to modulate the impact of pH on some properties however. Masking charges (120 mM salt) did lead to larger particles than when low salt was used under same pH. This was evident when pH was distant from pI (pH 7 to 9) where enzymes were not neutral globally, thus subject to repulsion. For pH 7 to 9, salt concentration was a determining factor for denaturation of secondary structures. Under a few particular set of conditions (pH, salt concentration), it was not possible to achieve denaturation of secondary structures (despite denaturation of tertiary structure at the “expected” T_m). The ability to retrieve a normal denaturation curve in the presence of limited amount of urea strongly suggests that aggregation or association is involved in the above mentioned phenomenon. Such phenomenon suggests the existence of an intermediate with preserved secondary structures but where tertiary structure is denatured. This intermediate would be available to EstGtA2 in the presence of specific pH and salt concentration conditions (pH 7, 20 mM salt; pH 8, 50 mM salt and pH 9, 120 mM salt). It appears that as we move pH further from pI, more salt is required to permit access to this intermediate, which is dependent on a fine balance between ionization status and electrostatic repulsion.

Examination of an EstGtA2 tridimensional model allows for possible identification of key residues involved in determining the denaturation/aggregation paths favored by the protein. On the basis of this model (that is based on MGL H-257, PDB entry 3RM3, 89% identity, shown in Supplementary Materials Fig. S5) a total number of 60 potentially charged and mostly exposed residues are suggested in EstGtA2 (19 positively charged K and R residues with 9 H residues when exposed to acidic pH; 32 negatively charged D

and E residues). Approximately 24% of EstGtA2 total residue composition can be charged under benign pH conditions. Some of these charged residues are involved in salt bridges that are important for EstGtA2 thermal stability as previously reported [17]. At pH 5, assuming expected His side chain pKa values (6.1), EstGtA2 would have an almost equal amount of positively charged (28) and negative charged residues (32). Using side chain expected pKa values, the theoretical isoelectric point (pI) was estimated to 5.03 which is very close to the experimentally determined pI = 4.90 (Fig. S4, this study). This information suggests that under alkaline conditions, EstGtA2 would be negatively charged as opposed to under acidic conditions where EstGtA2 has a low net charge (pH 6) or a null net charge (pH 5). The continuous surface electrostatic potential on EstGtA2 under acidic, neutral and moderately alkaline pH conditions was calculated as shown in Figure 5. For all simulations, especially at low ionic strength. At pH 5 and low ionic strength, EstGtA2 is a zwitterion (experimental pI = 4.9) and therefore has a null net charge (equal proportion of positively and negatively charged residues) as shown by the close to equal mix of surface electrostatic potential (white, blue and red) at all ionic strength calculated. The distribution of charge residues appears to be highly polarised, with an obvious cluster of positive charge on one face of EstGtA2 and a large negative surface on the other side. This strongly suggests that interactions between positive and negative patches of neighboring proteins should favor aggregation. At neutral pH, EstGtA2 is predicted to become negatively charged as the pH moves away from its pI as shown by the high proportion of red surface (negative potential). As expected, when the ionic strength is increased under neutral and moderately alkaline conditions, the counter ions (sodium and chloride were used here for this simulation) neutralize the residue charges on EstGtA2. At neutral pH, moderate ionic strength (120 mM salt) neutralized the negative charges of EstGtA2 to the extent where the protein became almost neutral, mimicking the surface calculated at pH 5 (120 mM). At moderately alkaline pH (pH 8), the presence of 120 mM salt neutralizes the charges but to a lesser extent than at pH 7 as shown by the higher residual negative potential (red surface) on the protein surface shown on the left side, compared with pH 7. The calculations are supporting the view that ionic

strength has an impact at neutral pH because it can decrease charge potential and decrease repulsion, despite the dominance of negative charges.

Figure 5

Proposed model for heat induced EstGtA2 denaturation and aggregation

After heat-induced denaturation, access to three different denatured/aggregated states is suggested. These paths are tentatively identified as: amorphous aggregation, partially folded aggregation controlled by surface charges, and soluble denaturation. Despite the new intermediate path proposed here (partially folded aggregation), our study is compatible with the model proposed by Tyedmers *et al.* (2010) [26] as described in Figure 6. This earlier model was adapted by introducing an intermediate stage, where tertiary structure is compromised but secondary structures are preserved. This associated state would be characterized by a size of 300 nm that we observed when secondary structure resisted denaturation despite heating up to 95°C. Our study suggest that access to paths A, B or C is not exclusive, i.e. that while most of EstGtA2 population may aggregate through path A, there may be another minor protein population favoring path B for example.

Although the EstGtA2 thermal denaturation pathways clearly distinguished under acidic and alkaline conditions, at neutral pH, a particular phenomenon was observed which is characterized by intermediate size particles ($R_h = 300$ nm). At pH 7, a nearly equal mix of small and medium size aggregates ($R_h = 8$ nm and 300 nm) was detected in addition to a small proportion of the large particles observed under acidic conditions ($R_h = 2000$ nm). The unique type of intermediate size particles ($R_h = 300$ nm) is a feature characteristic to neutral and moderately alkaline pHs and seems to protect the secondary structures from melting as observed by the absence of Far-UV CD transition upon heating, and this up to 95°C. Despite that no fusion curve may be obtained at pH 7, when looking at secondary structures (CD signal at 222 nm) due to the apparent high rigidity of the

alpha-helices, the tertiary structure actually unfolded as seen by near-UV CD (signal at 272 nm).

Figure 6

There is an abundant literature on the importance of beta sheet structures in the formation of organised aggregates such as amyloids [27,28]. Fändrich (2012) reported on a protofibrillar intermediate 10 x 400 nm in size, dominated by β structures that might correspond to the aggregation intermediate with 300 nm HR observed here. We did not detect any increase in beta sheet content (and no significant change in helix content either), but we cannot rule out that the native strands of EstGtA2 took part in promoting fibrils like the ones observed earlier. Hugonin et al. (2008) discussed an aggregation intermediate where alpha-helices but no β sheets were present. Here our results point to the formation of an α/β partially folded intermediate of moderate size (close to protofibrillar intermediate mentioned above).

Implications for industrial applications

As mentioned by Kumar et al. (2000), surface charged amino acids and salt bridges are known to be more abundant in thermophilic protein and to enhance their thermostability [19]. In the case of EstGtA2, we hereby demonstrated that they directly control the aggregation process during thermal denaturation. Moreover, the partially folded intermediate observed could be related to a protection mechanism in *Geobacillus thermodenitrificans*, enabling the protection and the refolding of enzymes after facing extreme temperature conditions. The understanding of the particular behavior of EstGtA2 could be relevant for determining its limitations and optimal industrial uses too, or shed light on strategies that may be considered to adapt it to harsh conditions for possible industrial applications. For low pH applications, EstGtA2 has a T_m that should prevent denaturation for applications up to 65°C. However, prolonged heating at higher temperature will result in irreversible loss of enzymatic activity. Under such conditions, protein engineers may have to consider modifying pH responsive exposed residues in the

hope to fight aggregation. At pH 7, T_m is also very high (74°C) and at low salt concentration, EstGtA2 will stand prolonged heating without activity loss. The same is true for pH 8 but at higher salt concentrations. At higher pH, EstGtA2 conformational stability decreases steeply and although it is not aggregation prone, its efficiency above 40°C is compromised. Under such conditions, aggregation is not an issue and conformational stabilisation via rational mutagenesis would help. For example, Charbonneau et al. (2013) did increase melting temperature of EstGtA2 at pH 10 by a His-Arg mutation in position 222 (H222R), which tends to confirm our interpretation [17].

Considering the impact of pH in determining behavior of EstGtA2, one may start by looking at possible salt bridges that would be formed or broken over the pH 7 to 9 range. The intuitive hypothesis would identify histidine residues because of their theoretical side chain pKa (6.04). Based on a homology model of EstGtA2 recently proposed by Charbonneau et al. [17], it appears that there would be 3 salt bridges at the surface of EstGtA2 involving histidine: E78-H110; D148-H197 and D194-H222. These histidines could be responsible of the modulation of aggregation paths upon pH variation and this can be verified using mutagenesis and characterisation of mutants. Unfortunately, the exact pKa of amino acid residues in a folded protein remains elusive. Further, without a reliable experimentally derived structural model (obtained from x-ray diffraction data or NMR) the identification of salt bridges remains approximate [29, 30]. We continue our efforts (collaboration with Prof. A. Berghuis, McGill University) in trying to elucidate EstGtA2 carboxylesterase. Crystallisation conditions are currently screened in the hope of generating stable crystals for a series of EstGtA2 mutants and identify key residues in controlling aggregation.

Conflict of interest

The authors declared there is no conflict of interest.

Acknowledgements

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (number RGPIN 138654). J.K.M. acknowledges the support of PROTEO scholarships. D.M.C acknowledges the help of Maximilian C. C. J. C. Ebert for access to Adaptive Poisson-Boltzmann Solver (APBS).

Contribution of the authors

M.B, F.M-M and J.K.M designed the study. J.K.M. performed all the experiments and wrote the initial draft of the manuscript. D.M.C performed the tridimensional structure analyses and the mapping of electrostatic potentials on EstGtA2 model. D.M.C organized the presentation of the data and wrote the final version of the manuscript which was helped by F.M-M and M.B. All authors contributed to the interpretation of the data and approved the final manuscript before publication.

Conclusion

The understanding of the different mechanisms occurring during thermal denaturation is crucial for developing better adapted industrial enzymes. Moreover, controlling irreversible aggregation is crucial for the recyclability of enzymes in a given industrial process where temperature changes are unavoidable. Here we reported 3 different denaturation/aggregation patterns observed during EstGtA2 heat-induced denaturation allowing for a better understanding of EstGtA2 denaturation behavior. At pH 7 and 8, depending on salt concentration, a folded but non-native species resisted secondary structure unfolding up to 100°C. Such folded intermediate was not accessible under acidic or alkaline pH. Ionic interactions at the surface of the protein are responsible for the pH and ionic strength dependent behavior of EstGtA2 and appear to control the aggregation path chosen by the enzyme.

References

1. Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F et al. (1992) The alpha/beta hydrolase fold. *Protein Eng* 5(3): 197-211. Doi: 10.1093/protein/5.3.197.
2. Bornscheuer UT (2002) Microbial carboxylesterases: classification, properties and applications in biocatalysis. *FEMS Microbiol Rev* 26: 73-81.
3. Fojan P, Fojan P, Jonson PH Petersen MTN, Petersen SB (2000) What distinguishes an esterase from a lipase: A novel structural approach. *Biochimie* (82): 1033-1041.
4. Arpigny JL, Jaeger K-E (1999) Bacterial lipolytic enzymes: classification and properties. *Biochem J* 343: 177-183.
5. Lenfant N, Hotelier T, Velluet E, Bourne Y, Marchot P, and Chatonnet A (2013) ESTHER, the database of the alpha/beta-hydrolase fold superfamily of proteins: tools to explore diversity of functions, *Nucleic Acids Res* 41, D423-429.
6. Fisher M, Pleiss J (2003) The lipase engineering database: a navigation and analysis tool for protein families. *Nucleic Acids Res* (31): 319-312.
7. Shao H, Xu L, Yan Y (2014) Biochemical characterization of a carboxylesterase from the archaeon *Pyrobaculum* sp. 1860 and a rational explanation of its substrate specificity and thermostability. *Int J Mol Sci* (15): 16885-16910.
8. Van den Burg B (2003) Extremophiles as a source for novel enzymes. *Curr Opin Microbiol* 6: 213-218.
9. Li X, Qian P, Wu SG, Yu HY (2014) Characterization of an organic solvent-tolerant lipase from *Idiomarina* sp. W33 and its application for biodiesel production using *Jatropha* oil. *Extremophiles* 18: 171-178.
10. Roberts CJ (2014) Therapeutic protein aggregation: Mechanism, design, and control. *Trends Biotechnol* 32(7): 372-380.
11. Fink AL (1998) Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Folding and design* 3: 923.
12. Tsolis AC, Papandreou NC, Iconomidou VA, Hamodrakas SJ (2012) A consensus method for the prediction of 'aggregation-prone' peptides in globular proteins. *Plos One* 8(1): e54175.

13. Tsai M-Y, Zheng W, Balamurugan D, Schafer NP, Kim BL, Cheung MS, Wolynes P G (2015) Electrostatics, structure prediction, and the energy landscapes for protein folding and binding. *Protein Sci* 25, 255-269.
14. Ahmad S, and Rao NM (2009) Thermally denatured state determines refolding in lipase: mutational analysis, *Protein Sci* 18, 1183-1196.
15. Majhi PR, Ganta RR, Vanam RP, Seyrek E, Giger K, Dubin PL (2006) Electrostatically driven protein aggregation: β -Lactoglobulin at low ionic strength. *Langmuir* 22(22): 9150-9159.
16. Augustyniak W, Brzezinska AA, Pijning T, Wienk H, Boelens R, Dijkstra BW, Reetz MT (2012) Biophysical characterization of mutants of *Bacillus subtilis* Lipase A evolved for thermostability: Factors contributing to increased activity retention. *Prot Sci* 21: 487-497.
17. Charbonneau DM, Beauregard M (2013) Role of key salt bridges in Thermostability of *G. thermodenitrificans* EstGtA2: Distinctive Patterns within the new bacterial lipolytic Enzyme subfamily XIII.2. *PLoS ONE* 8(10): e76675.
18. Charbonneau, D. M., and Beauregard, M. (2015) Correction: Role of Key Salt Bridges in Thermostability of *G. thermodenitrificans* EstGtA2: Distinctive Patterns within the New Bacterial Lipolytic Enzyme Family XV, *PLoS One* 10, e0136940.
19. Kumar S, Tsai C-J, Nussinov R (2000) Factors enhancing thermostability. *Protein Engineering* 13(3): 179-191.
20. Vetriani C, Maeder DL, Tolliday N, Yip KS, Stillman TJ et al. (1998) Protein stability above 100 degrees C: a key role for ionic interactions. *Proc Natl Acad Sci USA* 95: 12300-12305.
21. Knochel T, Pappenberger A, Jansonius JN, Kirschner K (2002) The crystal structure of indoleglycerol-phosphate synthase from *Thermotoga maritima*. Kinetic stabilization by salt bridges. *J Biol Chem* 277: 8626-8834.
22. Pace CN, Shirley BA, Thompson JA (1989) Measuring the conformational stability of a protein. In: TE Creighton. *Protein structure- A practical approach*. Oxford University Press. pp. 311-330.
23. Dolinsky TJ, Czodrowski P, Li H, Nielsen JE, Jensen JH, Klebe G and Baker NA (2007) PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations, *Nucleic Acids Res* 35, W522-525.

24. Baker NA, Sept D, Joseph S, Holst MJ, and McCammon JA (2001) Electrostatics of nanosystems: application to microtubules and the ribosome, *Proc Natl Acad Sci U S A* 98, 10037-10041.
25. Kelly SM and Price NC (2000) The use of circular dichroism in the investigation of protein structure and function, *Curr Protein Pept Sci* 1, 349-384.
26. Tyedmers J, Mogk A, Bukau B (2010) Cellular strategies for controlling protein aggregation. *Nature reviews Molecular Cell Biology* 11: 777-788.
27. Fändrich M (2012) Oligomeric intermediates in amyloid formation: structure determination and mechanisms of toxicity. *J Mol Biol* 421(4-5): 427-40.
28. Hugonin L, Barth A, Gräslund A, Perálvarez-Marín A (2008) Secondary structure transitions and aggregation induced in dynorphin neuropeptides by the detergent sodium dodecyl sulfate. *Biochim Biophys Acta* 1778(11): 2580-2587.
29. Yang A-S, Gunner MR, Sampogna R, Sharp K, Honing B (1993) On the calculation of pKas in proteins. *Proteins: Structure, Function and Bioinformatics* 15(3): 252-265.
30. Anderson DE, Becktel WJ, Dahlquist FW (1990) pH-induced denaturation of proteins: a single salt bridge contributes 3-5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry* 29(9): 2403-2408.

Tables

Table 1. Effect of pH and salt concentrations on EstGtA2 denaturation and aggregation. Apparent melting temperatures derived from CD data and visual appearance of protein solutions after heating to 95°C are presented for different pH and ionic strength. Aggregated species sizes (R_h (nm)) in EstGtA2 solutions were calculated by DLS after thermal runs and are indicated as superscripts.

pH	Sodium salt								
	20 mM			50 mM			120 mM		
	T_m (°C)	T_m^3 (°C)	Appearance	T_m (°C)	T_m^3 (°C)	Appearance	T_m (°C)	T_m^3 (°C)	Appearance
5	70.8 ± 0.7	n.d.	A	70.8 ± 0.2	n.d.	A	70.6 ± 0.1	n.d.	A
6	74.0 ± 0.15	66.3 ± 2	A ²⁰⁰⁰	74.2 ± 0.1	n.d.	A	73.9 ± 0.1	66.5 ± 0.5	A ²⁰⁰⁰
7	-	72.3 ± 0.1	Cloudy ⁸⁻³⁰⁰	78.4 ± 2.7	n.d.	A	74.1 ± 0.3	68.1 ± 0.5	A ²⁰⁰⁰
8	68.8 ± 0.6	70.4 ± 1.1	Clear ⁸	-	n.d.	Clear	-	68.4 ± 0.5	Cloudy ³⁰⁰⁻²⁰⁰⁰
9	63.6 ± 0.42	63.9 ± 0.2	Clear	65.6 ± 2.9	n.d.	Clear	-	59.3 ± 0.03	Cloudy ¹⁰⁻³⁰⁰
10	51.9 ± 0.2	n.d.	Clear	49.7 ± 0.1	n.d.	Clear	46.5 ± 3	n.d.	Clear
11	41.1 ± 1.4	n.d.	Clear	37.0 ± 2.1	n.d.	Clear	41.6 ± 0.9	n.d.	Clear

n.d.: Not determined.

-: indicates inability to calculate T_m due to secondary structures stabilization.

A: indicates visible aggregation.

Superscript number: Hydrodynamic Radius (R_h) in nm determined by DLS.

T_m : indicates thermal melting temperature of the secondary structure measured by Far-UV CD in triplicates.

T_m^3 : indicates thermal melting temperature of the tertiary structure estimated from near-UV CD.

Note that the appearance of the protein solution after thermal denaturation of the CD was determined by observation with the naked eye through the quartz cell.

Table 2. Comparison of EstGtA2 melting temperatures in the presence of urea (2 M). Melting temperatures are presented as function of pH (6, 7 and 8), without (0M) and with urea (2 M). Salt content was 20 mM. No transition or denaturation was observed at pH 7 and 0 M urea.

pH	<i>T_m</i> (°C)	
	Urea (0 M)	Urea (2 M)
6	74	68.4
7	-	66.7
8	68.7	62.1

Legends to figures

Fig 1. Secondary structures denaturation temperatures for EstGtA2 under various conditions. Apparent melting temperatures (derived from CD) are represented as mean of triplicate with standard error as function of pH in buffer A (20 mM), buffer B (50 mM) and buffer C (120 mM total sodium salt). Conditions leading to absence of denaturation transitions despite heating at 95°C are indicated by a star.

Fig 2. Conformational denaturation analysis by far-UV circular dichroism. CD spectra recorded at various temperatures during thermal denaturation in 20 mM salt at selected pH values. Typical spectra recorded at 25 and 100°C are shown as normal line and dashed lines respectively. CD spectra of samples treated at pH 5 (green line), 7 (black line) and 8 (red lines) are shown. Signal is expressed as molar CD.

Fig 3. Effect of pH and salt concentration on EstGtA2 hydrolysis of p-NPC8. Enzymatic activity measurement are presented as mean of triplicate with standard error as function of pH in 20 mM buffer (plain bars) and buffer containing 120 mM total salt concentration (striped bars). Bar shade gets darker as pH increases (light bars for pH 6, dark bars for pH 8).

Fig 4. Effect of pH and salt on EstGtA2 activity recovery after heating. Enzymatic activity measurements were performed after boiling the enzyme as function of pH in 20, 50 and 120 mM buffer (pH 7 is presented in gray and pH 8 in black). No residual activity (0%) was detected after heating at pH 5 or 6 (associated with large aggregates formation).

Fig. 5. Effect of pH and salt on the surface electrostatic potential of EstGtA2 at high temperature. Surface electrostatic potential mapped on EstGtA2 at pH 5, 7 and 8 (at 80°C) in the absence of salt and in the presence 120 mM monovalent salt. The color scale for electrostatics extends from -3 (red) to +3 eV (blue) (kT/e).

Fig 6. Proposed model of EstGtA2 aggregation behavior. Model for heat-induced denaturation of EstGtA2 is represented according to conformational analysis by CD and resulting size aggregates measured by DLS. Depending on pH and salt concentration, heated EstGtA2 can undergo 3 denaturation paths; amorphous aggregation leading to 2000 m species (A), formation of a partially folded intermediate, as represented by the 300 nm species (B) and soluble denaturation leading to the presence of 8 nm particles (C).

Figures

Figure 1

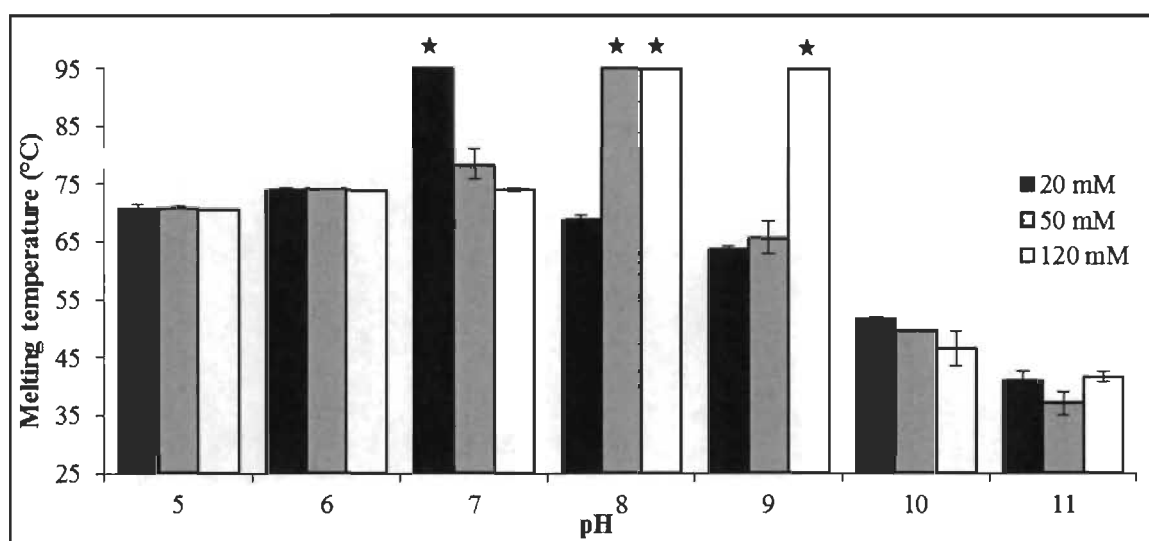


Figure 2

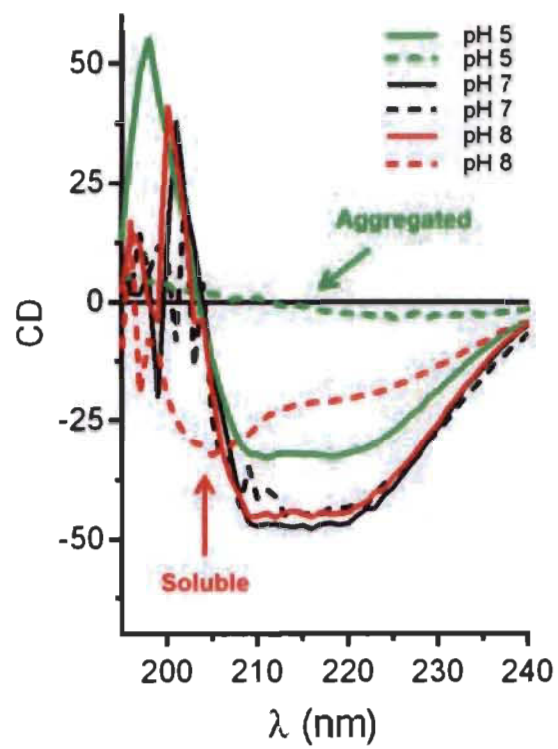


Figure 3

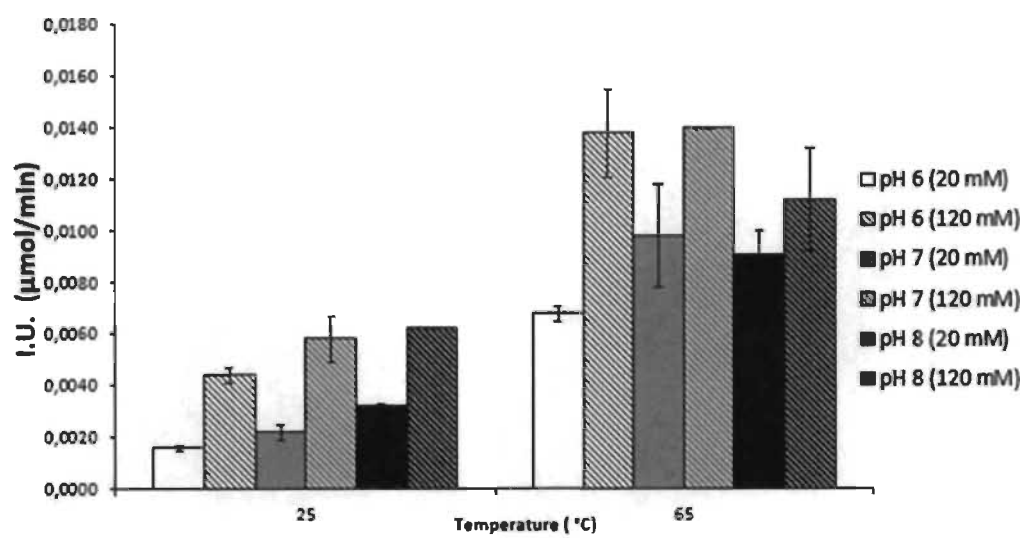


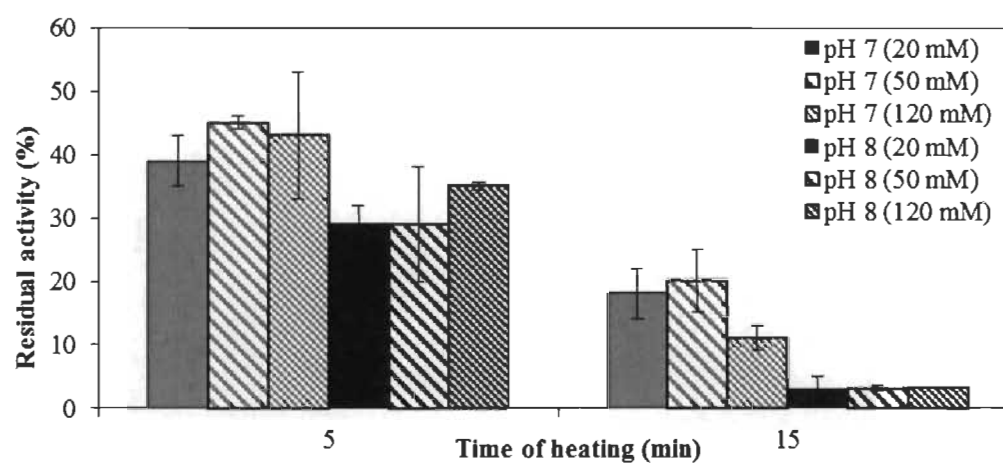
Figure 4

Figure 5

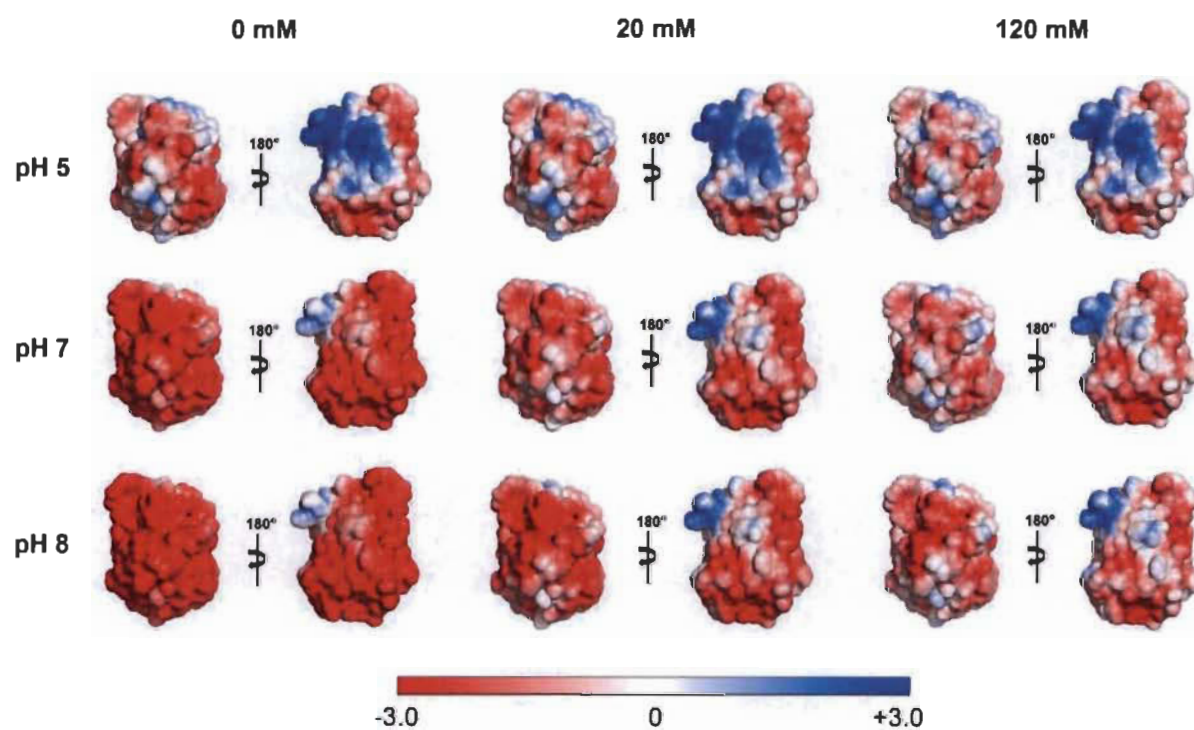
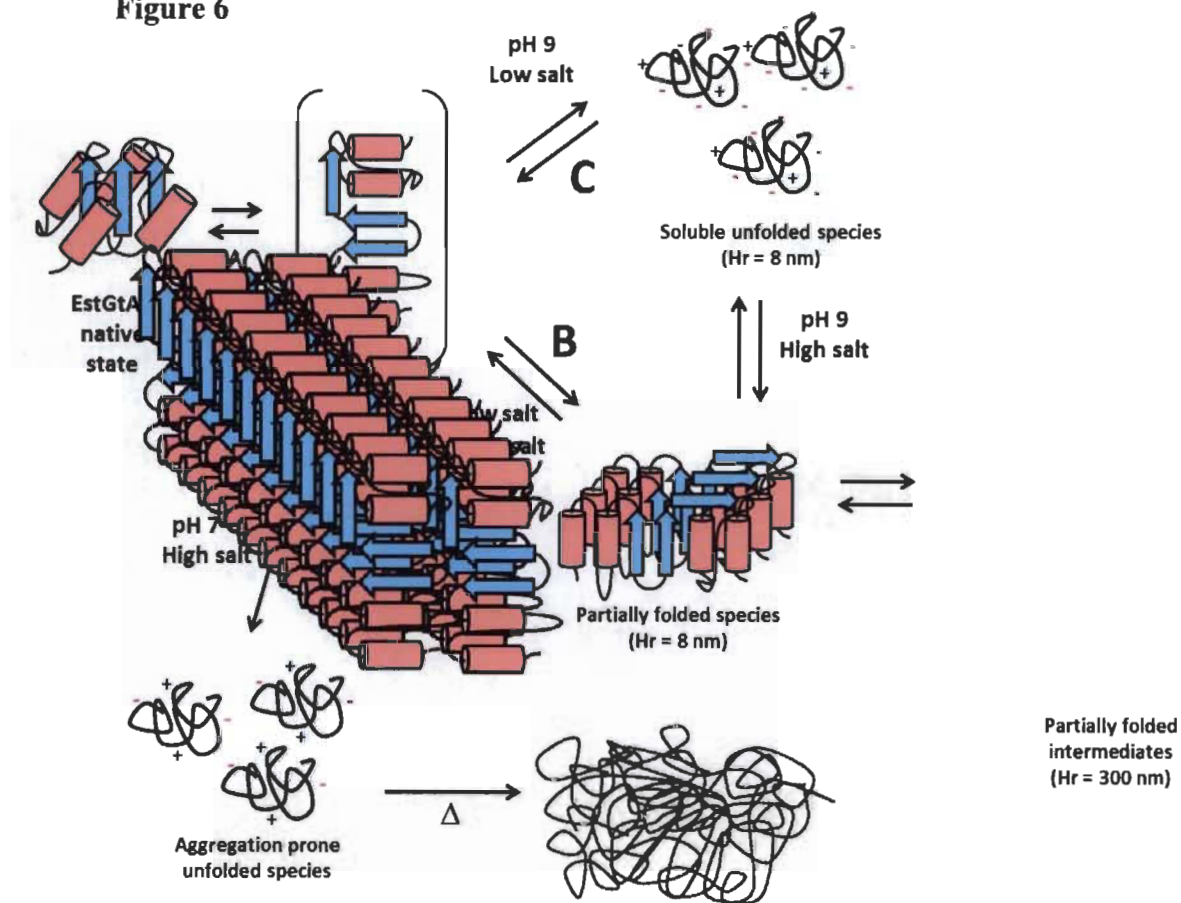
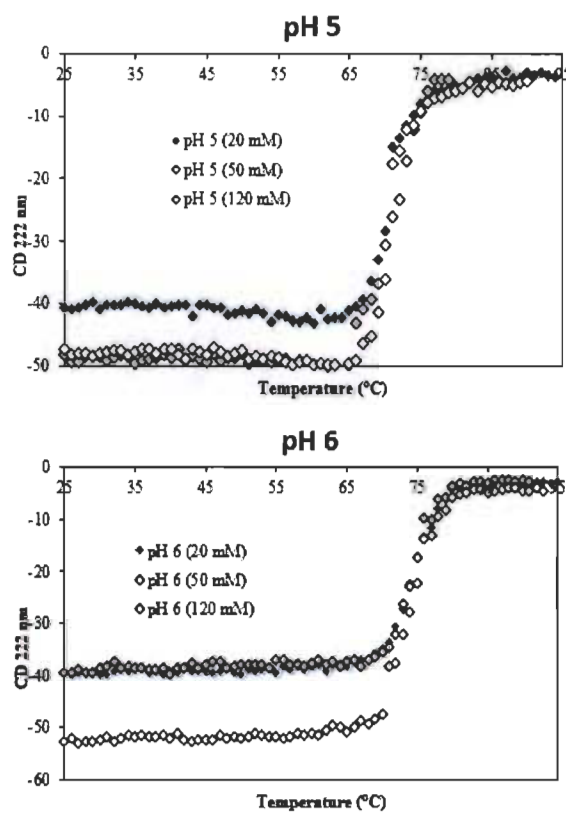


Figure 6



Supplementary data

Fig S1. Secondary structures thermal denaturation curves of EstGtA2 at different pH and ionic strength recorded by Far-UV CD at 222 nm (Helices).



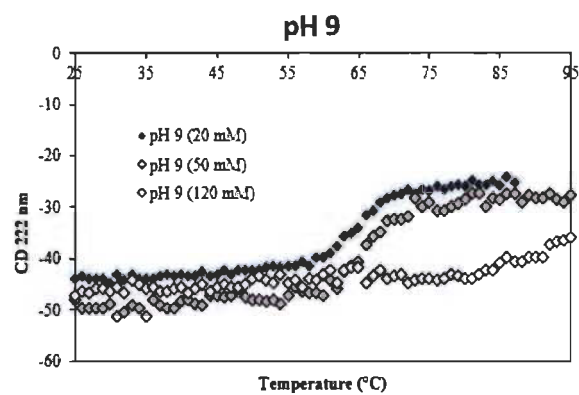
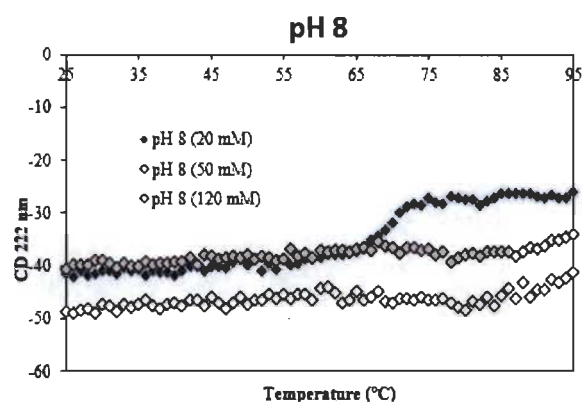
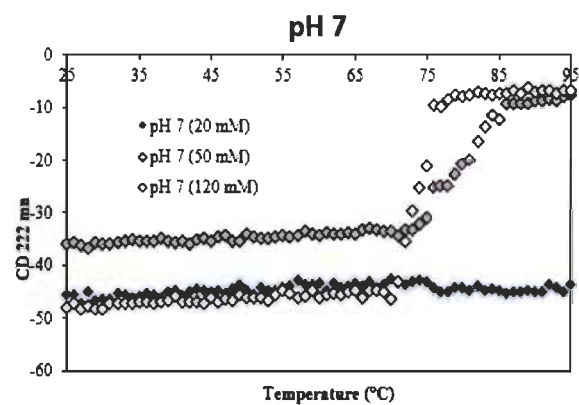


Fig. S2. Near-UV CD spectra of EstGtA2 recorded at 25 and 90°C at different pH and ionic strength.

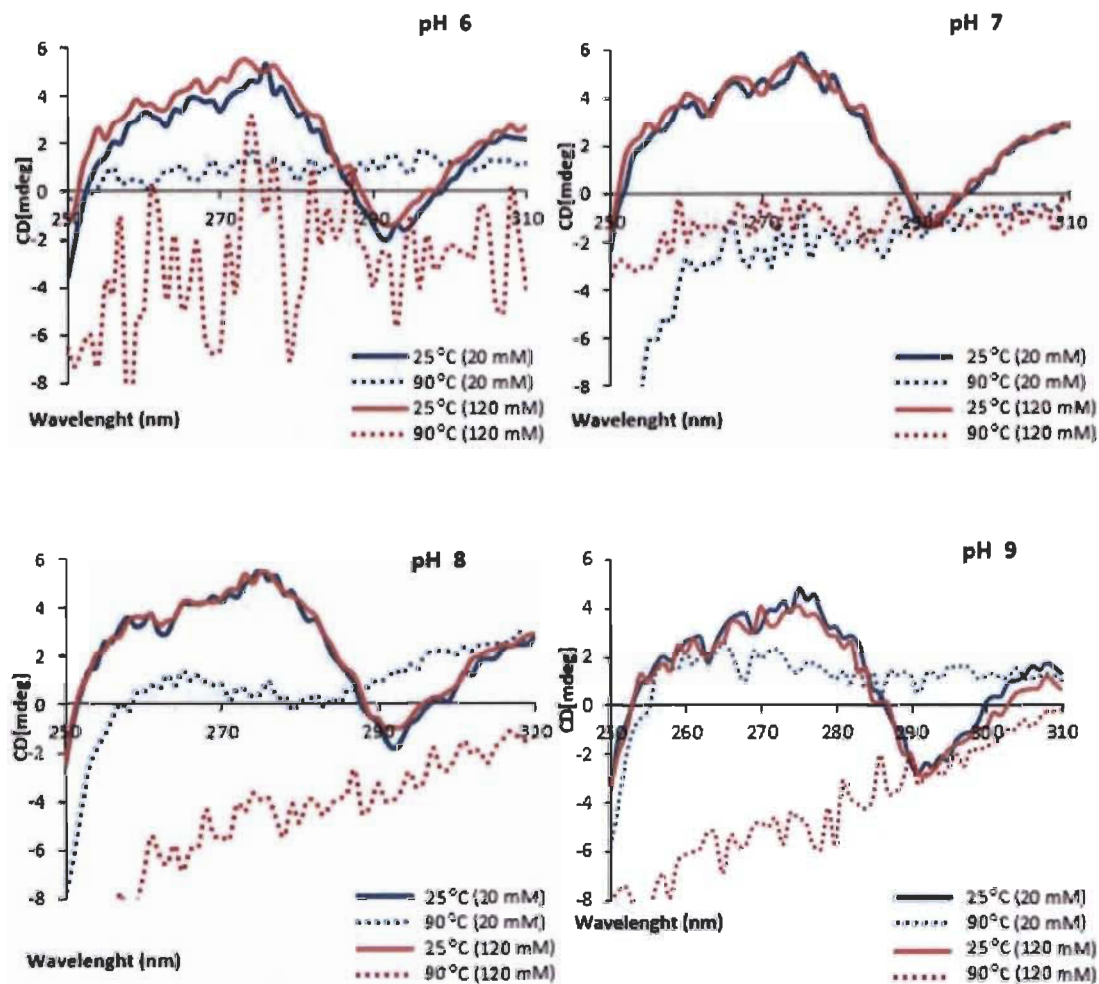


Fig S3. Photography of the CD quartz cell after heating at 100°C at pH 6 and pH 8. Aggregation was observed under acidic conditions while under alkaline conditions, an unfolded soluble state was observed.

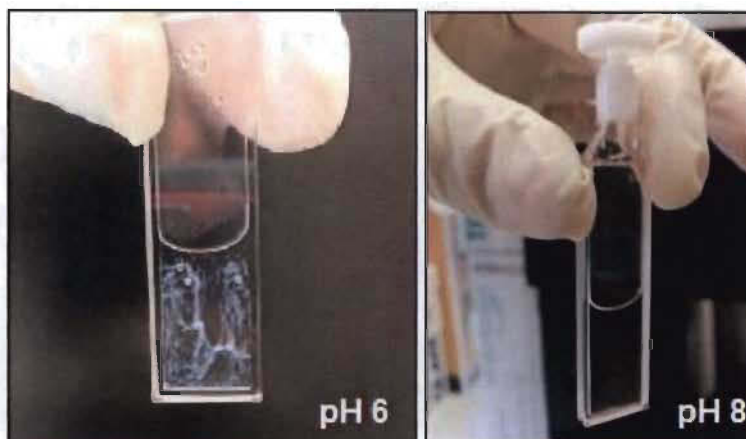


Fig. S4. Determination of EstGtA2 pI by isoelectric focusing. An isoelectric polyacrylamide gel is shown. 10 μ g (M1) and 5 μ g (M2) of marker protein standard (proteins with known pIs) were used to determine the isoelectric point of EstGtA2 (Est). Migration of bovine serum albumin protein is also shown (BSA; pI of 4.8).

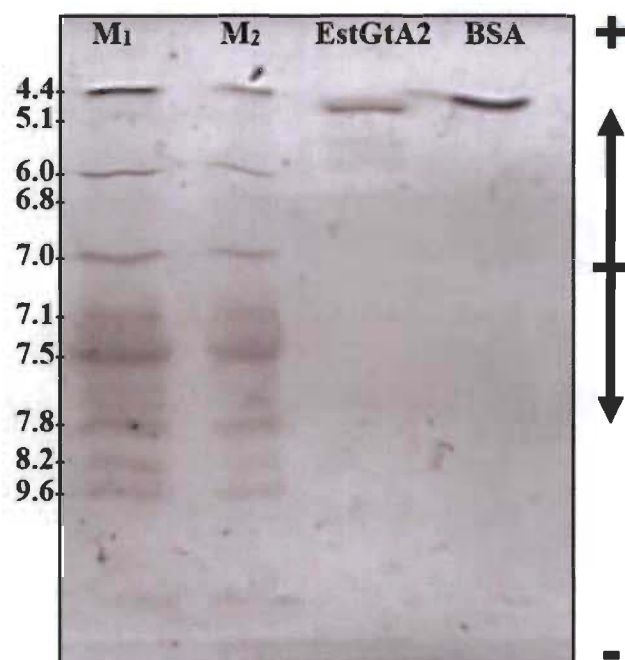
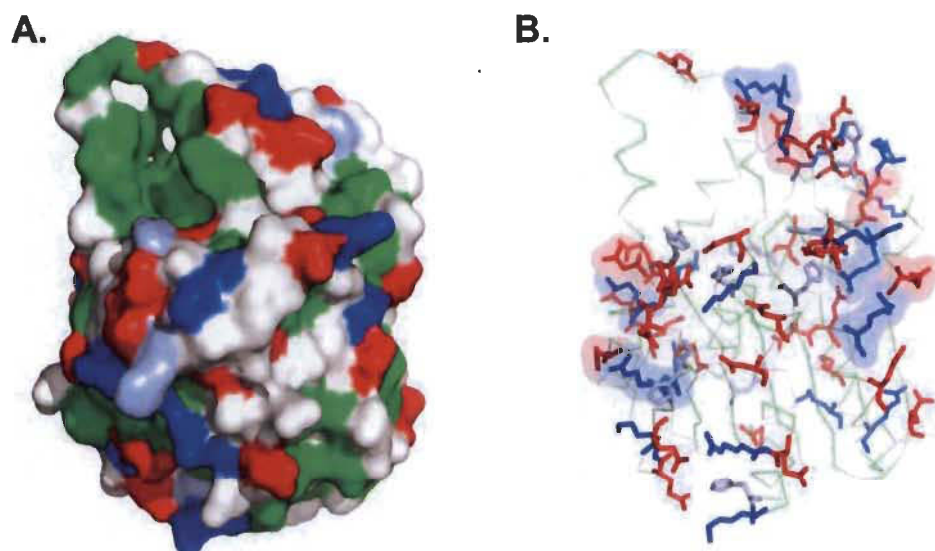


Fig. S5. Charged residues of EstGtA2. A. Surface view of an EstGtA2 model. Uncharged polar residues are shown in white, positively charged residues (R and K) in blue and negatively charged residues (D and E) in red. Histidine residues are shown in light blue. The hydrophobic residues of the core are shown in green. B. All 60 charged residues (predicted based on expected pKa values) of EstGtA2 are shown as sticks and colored as in A. The charged residues involved in salt bridges are shown with transparent surface.



CHAPITRE III

CONCLUSION

L'utilisation des enzymes comme biocatalyseurs dans les procédés industriels est un domaine innovant en biotechnologie. Elle permet notamment de diminuer l'énergie d'activation des réactions, donc diminuer l'énergie nécessaire pour générer certains produits, mais aussi de réduire la quantité d'agents chimiques utilisés et ainsi diminuer les déchets toxiques rejetés dans l'environnement. Parmi les enzymes bactériennes thermophiles identifiées et caractérisées, la plupart de celles-ci ne vont pas réussir de façon optimale la bioconversion désirée puisqu'elles ne sont pas totalement adaptées pour résister aux conditions particulières retrouvées en industrie (température extrême, force ionique, pH). La caractérisation de nouvelles enzymes thermophiles est donc nécessaire pour le développement des biocatalyseurs qui seront mieux adaptés aux conditions industrielles.

À ce jour, il nous est impossible de prédire avec certitude le comportement qu'adoptera une enzyme dans différentes conditions sans avoir à passer par une caractérisation exhaustive en laboratoire, ce qui représente une demande considérable en termes de charge de travail. Dans cet ordre d'idée, la compréhension des relations entre la séquence d'acides aminés et les propriétés enzymatiques est de mise pour l'utilisation, le recyclage et la mise au point de variantes d'enzymes. Cette compréhension permettrait de faire des prédictions (activité, stabilité et applications potentielles) lors de l'annotation génomique des séquences d'enzymes nouvellement identifiées. Elle permettrait aussi de générer des variantes d'enzymes (par ingénierie des protéines) ayant des propriétés désirables pour une application donnée. En effet, l'identification d'éléments structuraux en lien avec la séquence en acides aminés permettant de prédire les propriétés d'une enzyme permettra de faciliter la mise au point de ces biocatalyseurs et ainsi réduire le temps de développement nécessaire. La comparaison des éléments conservés au sein

d'une même famille d'enzyme est un indicateur des éléments structuraux importants pour les propriétés de ces enzymes. Par exemple, chez les enzymes thermostables, il a été démontré que certains facteurs sont responsables de la stabilité à haute température. Les ponts salins sont entre autres reconnus pour être plus présents chez les enzymes thermostables et c'est sur ce facteur que les études sur la carboxylestérase EstGtA2 ont été basées.

EstGtA2 est une nouvelle enzyme lipolytique, de type carboxylestérase, qui a été isolée de la souche thermophile *Geobacillus thermodenitrificans*. L'enzyme recombinante a été produite et caractérisée pour la première fois dans notre laboratoire (Charbonneau 2010, 2012a, 2012b, 2013). La caractérisation d'EstGtA2 a été réalisée par mon collègue David M. Charbonneau, ce qui a d'abord permis de mieux comprendre les facteurs influençant la thermostabilité de l'enzyme. Il a été démontré que la thermostabilité d'EstGtA2 variait en fonction du pH et que les ponts salins présents à la surface de l'enzyme étaient directement impliqués dans la stabilité thermique de l'enzyme. Un modèle de la structure tridimensionnelle d'EstGtA2 a été généré à partir d'une protéine homologue cristallisée (89 % d'identité), permettant d'identifier les ponts salins conservés au sein de la famille de l'enzyme. Le rôle de ces ponts a été étudié par mutagenèse dirigée (Charbonneau et Beauregard, 2013).

Les résultats présentés dans le deuxième chapitre de ce document sont la continuité de la caractérisation d'EstGtA2 dont le but est d'identifier des éléments de structures clés pour la thermostabilité de la carboxylestérase EstGtA2. Pour arriver à comprendre l'impact d'une variation de pH et de force ionique sur sa thermostabilité et préparer l'enzyme pour de futures applications industrielles, l'enzyme a été produite de façon recombinante et purifiée en laboratoire. Différents tests de dénaturation thermique ont été réalisés sous différentes conditions (pH et forces ioniques) dans le but d'étudier la stabilité des structures secondaires et tertiaires de l'enzyme. La morphologie et la taille des agrégats après dénaturation ont aussi été étudiées dans le but de comprendre le type d'agrégation et/ou d'association de l'enzyme lors de sa dénaturation. De plus, des tests biochimiques ont été réalisés dans le but de vérifier l'impact de la force ionique et du pH

sur l'activité catalytique de l'enzyme. Des tests de thermostabilité ont aussi été faits dans le but d'étudier la capacité de l'enzyme à se replier et à retrouver son activité après chauffage. Finalement, des essais de cristallisation d'EstGtA2 ont été faits en collaboration avec le Dr Albert Berghuis de l'Université McGill dans le but d'élucider la structure tridimensionnelle de l'enzyme par diffraction de rayon X.

3.1 Synthèse et interprétation des résultats obtenus (sommaire)

3.1.1 Analyse conformationnelle des structures secondaires

Les dénaturations thermiques réalisées à différents pH (6-11) et force ioniques (20, 50 et 120 mM) avaient pour but de nous aider à comprendre le rôle de ces paramètres sur la thermostabilité d'EstGtA2. Tout d'abord les températures de fusion des structures secondaires (T_m) ont été comparées (voir tableau 1, chapitre II). Il a été observé que le pH jouait un rôle considérable au niveau de la thermostabilité de l'enzyme. Les analyses des températures de fusion en fonction du pH ont montré une meilleure thermostabilité à pH neutre (courbes montrant une forme de cloche). Les T_m diminuaient en s'éloignant du pH optimal. Contre toute attente, la variation de force ionique n'a pas eu un impact à tous les pH. En condition de pH acide (pH 5 et 6), aucun impact majeur n'a été observé. Les mêmes résultats ont été observés en condition de pH basique (pH 10-11), bien que les T_m fussent moins élevés qu'en pH acide.

C'est sous des conditions de pH neutre à légèrement alcalin (pH 7-9) que l'effet de la variation de force ionique s'est révélée. À pH 7 (20 mM), aucune perte de structure secondaire n'a été observée, même après chauffage jusqu'à 95 °C, ce qui constitue un phénomène très inattendu, surtout lorsqu'on arrive à dénaturer les structures secondaires au pH voisin (6 et 8) à des températures de 74 et 69 °C respectivement. Lorsque la force ionique a été augmentée à 50 mM, le même phénomène (conservation des structures secondaires au-delà de 95 °C) s'est reproduit à pH 8. Une fois la force ionique augmentée à 120 mM, le phénomène s'est aussi produit à pH 8 et 9. Ces résultats suggèrent que le

pH est critique pour la dénaturation des structures secondaires, mais que la variation de force ionique joue un rôle limité, se situant autour des pH 7 à 9.

L'apparence des solutions enzymatiques après dénaturation a aussi été inspectée. À faible force ionique (20 mM), les dénaturations ont mené à la formation de larges particules insolubles pour les pH 5 et 6, comparativement aux conditions alcalines qui n'ont présenté aucune turbidité (pH 8-11). À pH 7 (20 mM), un niveau intermédiaire de turbidité a été observé en contraste avec les autres pH. Le même état turbide a été observé aux conditions pour lesquelles la conservation des structures secondaires à 95 °C a été notée (pH 8 et pH 9 à 120 mM). Ces résultats démontrent qu'une variation de force ionique permettrait de contrôler la stabilité thermique de l'enzyme et le type d'agrégation obtenu. En d'autres mots, quand les ponts salins de surface sont affectés par les charges en solution, on voit une modification du type d'association pendant la dénaturation.

3.1.1 Analyse conformationnelle des structures tertiaires

Dans le but d'étudier la stabilité thermique de la structure tertiaire (vérifier la sensibilité de l'organisation globulaire et de la compacité de l'enzyme au chauffage), des tests de dénaturation thermiques ont été réalisés. Les températures de fusion des structures tertiaires (T_m^3) ont été calculées et comparées entre les différentes conditions de pH et de forces ioniques (voir tableau 1, chapitre II). Les résultats obtenus à pH 7 (20 mM) (conditions menant à une protection des structures secondaires) démontrent que l'enzyme a perdu sa compacité (les acides aminés aromatiques ne sont plus contraints), bien que les structures secondaires soient conservées dans ces conditions. En général, à tous les pH, la stabilité des structures tertiaires n'est que très peu influencée par le pH ou la force ionique, vu les très faibles variations entre les résultats (résultats variant de 66 à 74 °C).

3.1.2 Analyse de l'agrégation d'EstGtA2

Des tests de diffusion dynamique de la lumière pendant les dénaturations thermiques ont été réalisés dans le but d'étudier les types d'agrégats et/ou d'association obtenus et d'évaluer leur taille. Cette méthode, qui est basée sur le mouvement brownien des particules en suspension, permet de déterminer le rayon hydrodynamique des particules présent dans l'échantillon. Les rayons hydrodynamiques ont été étudiés à trois différents pH (6,7 et 8) et à deux concentrations en sel (20 et 120 mM). En dépit de la force ionique utilisée, des agrégats de 2000 nm ont été détectés en condition de pH acide (pH 5-6), tandis que des tailles de 8 nm ont été observées pour les conditions de pH alcalin (pH 10-11) (Tableau 3.1). À pH 7, un troisième type de particule a été identifié, présentant un rayon hydrodynamique de 300 nm. Plus précisément, on note un mélange de particules (8 – 300 nm) lorsque la force ionique est faible (20 mM) et un mélange de plus grosses particules (300 – 2000 nm) lorsque la force ionique est élevée (120 mM). À pH 8 et 9, les mêmes effets sont observés, mais à force ionique plus élevée, c'est-à-dire que lorsque le pH est augmenté, il faut augmenter la force ionique afin d'atteindre l'état de dénaturation intermédiaire (là où la solution présente des particules de 300 nm). Les analyses de diffusion dynamique de la lumière sont en accord les inspections visuelles des solutions et suggèrent donc la formation de trois types d'agrégats et/ou d'association suite à la dénaturation thermique d'EstGtA2 à différents pH et forces ioniques, soit un état d'agrégation irréversible (2000 nm), un état soluble et réversible (8 nm) ainsi qu'un état intermédiaire (300 nm) correspondant à la protection des structures secondaires à haute température.

3.1.3 Impact du point isoélectrique d'EstGtA2

Dans le but de voir si l'agrégation était due au fait que la protéine était à son pI (menant à une charge moléculaire neutre), la focalisation isoélectrique a été utilisée pour déterminer le point isoélectrique d'EstGtA2. On se rappelle que les protéines ont tendance à s'agréger lorsqu'elles sont à un pH correspondant à leur point isoélectrique puisque leur charge globale est neutre (il n'y a pas de répulsion entre les protéines). Un point

isoélectrique de 4.9 a été déterminé expérimentalement pour EstGtA2. Ce résultat peut expliquer l'agrégation de l'enzyme à pH 5 et 6 puisqu'elle est globalement neutre à pH 4.9. De même, ce pI est compatible avec l'absence de gros agrégats à pH alcalin.

3.1.4 Effet du pH et de la force ionique sur l'activité enzymatique

Des tests d'activité enzymatique avec un substrat commercial, le pNP-C8, ont été effectués à 25 et 65 °C. Comme le produit est coloré, il est facile de quantifier la quantité de produit formé par mesure d'absorption et ainsi calculer avec le coefficient d'extinction molaire la vitesse initiale de la réaction ($\mu\text{mol}/\text{min}$). Pour ce test, une plage de pH autour du pH neutre a été choisie (pH 6, 7 et 8). Pour les deux températures testées (22 et 65 °C), une augmentation de la vitesse de réaction lorsque la force ionique est augmentée a été observée. EstGtA2 était plus active en présence de force ionique de 120 mM et à 65 °C. Ces résultats ne permettent aucune interprétation en lien avec les états d'agrégation et d'association aux différents pH. Il faut noter que les températures testées ici ne permettent pas de dénaturer EstGtA2.

3.1.5 Effet du pH et de la force ionique sur la thermostabilité

En industrie, la température des bassins n'est pas stable, il faut que l'enzyme soit résistante aux variations de température. Si l'enzyme ne peut pas se replier après chauffage, elle n'est pas recyclable, ce qui augmente les coûts. Pour mesurer le repliement, l'enzyme a été chauffée à 100 °C pendant 5 et 15 minutes, puis refroidi à température pièce pendant 20 minutes, avant de mesurer l'activité résiduelle. À pH 6, l'enzyme n'avait aucune activité résiduelle, ce qui corrobore l'hypothèse des agrégats non solubles formés sous pH acide (pH 5-6). À pH 8, l'enzyme montre une activité résiduelle, ce qui coïncide avec le fait que l'agrégation est soluble ou sous une forme intermédiaire. Étonnamment, c'est à pH 7 qu'a été observée la meilleure activité résiduelle (pH où les agrégats sont de tailles variables, allant de 300 nm à 2000 nm en fonction de la force ionique. Le fait qu'on trouve de l'activité après 5 min pour les conditions où l'on trouve des agrégats non solubles peut s'expliquer par le fait que ce ne sont pas toutes les enzymes qui ont eu le

temps de s'agréger en 15 minutes. Ce temps ne se compare pas au temps de 100 minutes utilisées pour enregistrer la dénaturation thermique (CD, DLS). Ceci laisse supposer qu'il reste une fraction d'enzymes qui n'est pas totalement agrégée de façon non soluble, ce qui concorde aussi avec les résultats de diffusion dynamique de la lumière. Il faut savoir que la thermostabilité est dépendante de plusieurs facteurs comme le T_m , les structures secondaires et tertiaires ainsi que le type d'association.

3.1.6 Sommaire des analyses

Il a été démontré que le comportement agrégatif de l'enzyme face à une augmentation de température est dépendant du pH ainsi que de la force ionique présente dans l'environnement de l'enzyme, mais dans une moindre mesure. Nous avons pu observer l'apparence de la solution (turbidité) de l'enzyme après dénaturation thermique et les observations ont pu être validées avec des mesures de rayon hydrodynamique par DLS. Nous avons été en mesure d'identifier une morphologie distincte (particules de 300 nm), qui semble marquer la transition entre la dénaturation soluble (pH alcalin) et non soluble (pH basique) de l'enzyme. Comme le pH contrôle le type d'agrégats obtenus de façon plus évidente que l'augmentation de force ionique (bien qu'elle joue un rôle marqué autour des pH neutres (7-9)), nous avons conclu que les résidus ionisables, répondant aux changements de pH, sont directement impliqués dans le mécanisme d'agrégation d'EstGtA2. Nous suggérons que les ponts ioniques se formant ou se déformant entre les pH 7 et 9 ont un impact direct sur la stabilité thermique d'EstGtA2 à pH alcalin.

Une étude par bio-informatique des acides aminés chargés à la surface d'EstGtA2, réalisée par mon collègue David M. Charbonneau, nous a permis d'identifier certains résidus ionisables pouvant être responsables de l'agrégation irréversible d'EstGtA2. Entre autres, les histidines sont les premiers acides aminés à prendre en considération dû à leur pKa de 6.01. Selon le modèle tridimensionnel d'EstGtA2 basé par homologie, on trouverait trois histidines formant des ponts salins à la surface d'EstGtA2 (E78-H110; D148-H197 and D194-H222).

3.2 Limitations

L'utilisation de modèles de structure de protéines basés sur l'homologie est une pratique courante en biochimie et en enzymologie industrielle. Pour ce faire, on utilise une protéine dont la séquence en acides aminés est similaire à la protéine d'intérêt et dont la structure tridimensionnelle a été résolue. Le problème avec les modèles par homologie, c'est qu'ils ne sont pas nécessairement représentatifs à 100 % de la véritable structure tridimensionnelle de l'enzyme d'intérêt. Notons qu'un changement d'un acide aminé dans la séquence pourrait changer de manière considérable une structure calculée par homologie. De ce fait, nul ne peut prédire sans risque le positionnement d'une chaîne latérale d'acide aminé donné, ni son implication dans un pont salin.

De plus, lorsque la présence de ponts salins est basée sur ce type de modèle, il faut se souvenir que les pKa des chaînes impliquées peuvent être changés par l'environnement. Une histidine, avec un pKa attendu de 6.01, peut en réalité avoir un pKa « environnemental » de 7.5 (ces chiffres ne sont donnés qu'à titre d'exemple et ne constitue en aucun cas des valeurs réelles). La méthode de mutagenèse dirigée, qui consiste à changer un acide aminé par un autre dans la séquence protéique, reste alors la seule option afin de valider le rôle d'un acide aminé dans une protéine. Rappelons cependant que le temps pour caractériser l'enzyme muté est extrêmement long et qu'on ne peut connaître l'impact sur la structure 3D de l'enzyme sauf si la cristallisation est réussie.

3.3 Perspectives de recherche

EstGtA2 représente un modèle intéressant pour l'étude des mécanismes d'agrégation des protéines globulaires pour plusieurs raisons : elle possède un repliement tridimensionnel commun à plusieurs protéines globulaires (α/β), elle est constituée d'une chaîne d'acides aminés de taille moyenne (249 acides aminés), elle montre des comportements uniques lors des dénaturations thermiques qui sont totalement

dépendantes du pH et de la force ionique de son environnement et finalement son activité enzymatique est facilement quantifiable avec des tests de spectrophotométrie.

L'obtention d'une structure tridimensionnelle par cristallographie de rayon x permettrait de valider la présence des ponts salins étudiés dans les études précédentes. Avec les structures de mutants, on pourrait être en mesure de comprendre l'impact de ses mutations sur le site actif. On pourrait mieux comprendre le comportement unique menant à la protection des structures secondaires. Or jusqu'à maintenant, nos efforts pour cristalliser EstGtA2 sont restés vains.

Il serait intéressant de tester l'impact de produits stabilisateurs, comme des ligands ou des détergents (substrats, Tween) sur la thermostabilité de la carboxylestérase (dénaturation thermique, type et taille des agrégats). Est-ce qu'on est capable d'inhiber l'agrégation de manière à ce que chaque unité se dénature sans s'associer avec les autres unités? De cette façon, on pourra permettre la renaturation de l'enzyme et augmenter la recyclabilité de l'enzyme.

Les études fondamentales portant sur les mécanismes d'agrégation des enzymes en solution aideront à comprendre les mécanismes impliqués. Une meilleure compréhension de ces mécanismes permettra de mieux cibler les modifications moléculaires qui stabiliseront les enzymes. Que ce soit dans un but médical ou industriel, une compréhension accrue de l'agrégation des protéines en général aura un impact majeur en science.

3.4 Conclusion

Les carboxylestérases sont des enzymes qui possèdent plusieurs caractéristiques intéressantes pour agir à titre de biocatalyseur dans les procédés industriels (spécificité, résistance). Cependant, plusieurs de ces enzymes ne sont pas aussi performantes que souhaité dû aux conditions particulières auxquelles elles doivent faire face pendant les processus industriels. L'identification et la caractérisation de nouvelles enzymes

permettront de développer des enzymes encore plus résistantes face aux conditions extrêmes rencontrées dans les bassins d'usines. La compréhension des différents mécanismes qui se produisent pendant la dénaturation thermique des enzymes ou face à différentes conditions de forces ioniques nous permettra de mieux comprendre la relation entre la séquence en acides aminés et les propriétés physico-chimiques des enzymes. De cette façon, il sera beaucoup plus évident de cibler les mutations favorables à la stabilité d'une enzyme pour un procédé en particulier.

Dans cette étude, nous avons étudié le comportement de la carboxylestérase EstGtA2 lors de dénaturations thermiques dans différentes conditions de pH et de forces ioniques. Les résultats nous ont permis d'établir un modèle d'association et d'agrégation de l'enzyme dépendamment des conditions étudiées et nous ont permis de mieux comprendre le comportement de l'enzyme face à ces conditions. L'impact du pH et de la force ionique démontre clairement le rôle des interactions électrostatiques au sein d'EstGtA2. Les interactions électrostatiques formées ou non aux différentes conditions de pH et de forces ioniques contrôlent le chemin de dénaturation pris par l'enzyme. Le pH et la force ionique déstabilisent les interactions ioniques à la surface de la protéine et ceci a un impact sur l'endroit où est envoyée l'enzyme dans le diagramme du paysage énergétique (« landscape » énergétique menant à la formation d'agrégats solubles ou non solubles). Au niveau de l'applicabilité de l'enzyme, ce critère est très important. Pour avoir le contrôle sur le chemin pris par l'enzyme, il faut identifier les résidus ionisables responsables de l'agrégation non soluble et les muter dans le but de stabiliser l'enzyme dans une conformation soluble. Si la carboxylestérase EstGtA2 se retrouve dans un milieu alcalin où la force ionique est élevée, on veut pouvoir se diriger vers un état stable, qui se replie après dénaturation, et non vers des agrégats non solubles, là où il n'y a pas de retour en arrière possible.

Selon les études basées sur le modèle par homologie, on a pu cibler trois résidus histidines formant des ponts salins à la surface de l'enzyme qui pourrait être responsables de ce type de comportement. L'élucidation de la structure 3D, dans un futur proche, nous permettra de cibler avec précision les interactions électrostatiques à étudier par mutagènes dirigés et ainsi identifier les éléments de structures clés dans la stabilité de l'enzyme.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Arpigny, J.L. et Jaeger, K.-E. (1999). Bacterial lipolytic enzymes: classification and properties. *Biochem J* 343: 177-183.
- Bornscheuer, U.T. (2002). Microbial carboxylesterases: classification, properties and applications in biocatalysis. *FEMS Microbiol Rev.* 26: 73-81.
- Bajpai D. et V.K. (2007). Laundry detergents: An overview. *Journal of oleo science.* 56(7): 327-340.
- Charbonneau, D.M. (2014). *Éléments structuraux essentiels à la thermostabilité de nouvelles enzymes lipolytiques*. Thèse de doctorat en biophysique et biologie cellulaires. Université du Québec à Trois-Rivières, Trois-Rivières.
- Charbonneau, D.M. et Beauregard, M. (2013). Role of key salt bridges in thermostability of *G. denitrificans* EstGtA2: Distinctive patterns within the new bacterial lipolytic enzyme family XV. *PLoS ONE* 8(10): e76675.
- Charbonneau, D.M., Meddeb-Mouelhi, F., Boissinot, M., Sirois, M. et Beauregard, M. (2012a). Identification of thermophilic bacterial strains producing thermotolerant hydrolytic enzymes from manure compost. *Indian J of Microbiol* 52(1): 41-47.
- Charbonneau, D.M., Meddeb-Mouelhi, F. et Beauregard, M. (2012b). N-terminal Purification Tag Alters Thermal Stability of the Carboxylesterase EstGtA2 from *G. thermodenitrificans* by Impairing Reversibility of Thermal Unfolding. *Protein and peptide letters* 19(3): 264-269.
- Charbonneau, D.M., Meddeb-Mouelhi, F. et Beauregard, M. (2010). A novel thermostable carboxylesterase from *Geobacillus thermodenitrificans*: Evidence for a new carboxylesterase family. *The journal of biochemistry* 148(3): 299-308.
- Charnock, S.J. et McCleary, B.V. (2005). Les enzymes : Applications industrielles et analytiques. *Revue des Œnologues* 116: 1-5.
- Dickerson, R.E. (2008). Max Perutz and the secret of life, by Georgina Ferry. *Protein science* 17(2): 377-379.
- Dill, K. A. (1997). From Levinthal to pathways to funnels. *Nat. Struct. Biol* 4: 10-19.

- Fink, A.L. (1998). Protein aggregation: folding aggregates, inclusion bodies and amyloid. *Folding and design* 3: 923.
- Fisher, M. et Pleiss, J. (2003). The lipase engineering database: a navigation and analysis tool for protein families. *Nucleic Acids Res* (31): 319-312.
- Fojan, P., Jonson, P.H., Petersen, M.T.N et Petersen, S.B. (2000). What distinguishes an esterase from a lipase: A novel structural approach. *Biochimie* 82(11): 1033-1041.
- Garrett, R.H. et Grisham, C.M., *Biochemistry* (2010). Brooks/Cole Cengage Learning, Boston, MA, ed. 4.
- Kirk, O., Borchert, T.V. et Fuglsang, C.C. (2002). Industrial enzymes applications. *Current opinion in biotechnology* 13: 345-351.
- Kourist, R., Brundiek, H. et Bornschuer, U.T. (2010). Protein engineering and discovery of lipases. *Eur J Lipid Sci Technol* 112: 64-74.
- Knochel, T., Pappenberger A., Jansonius, J.N. et Kirschner, K. (2002). The crystal structure of indoleglycerol-phosphate synthase from *Thermotoga maritima*. Kinetic stabilization by salt bridges. *J Biol Chem* 277: 8626-8834.
- Kumar, S., Tsai, C.-J. et Nussinov, R. (2000). Factors enhancing protein thermostability. *Prot Eng* 13: 179-191.
- Li, X., Qian, P., Wu, S.G. et Yu, H.Y. (2014). Characterization of an organic solvent-tolerant lipase from *Idiomarina* sp. W33 and its application for biodiesel production using Jatropha oil. *Extremophiles* 18: 171-178.
- Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolov, F. et al. (1992) The alpha/beta hydrolase fold. *Protein Eng* 5(3): 197-211. Doi: 10.1093/protein/5.3.197.
- Pauling, L. (1948). *The Nature of Chemical Bond*, 2nd Edition, Cornell University Press, Ithaca.
- Pauling, L. (1960). *The Nature of Chemical Bond*, 3rd Edition, Cornell University Press, Ithaca.
- Pauling, L. et Corey, R.B. (1951). "Atomic Coordinates and Structure Factors for Two Helical Configurations of Polypeptide Chains" *Proc. Natl. Acad. Sci. USA* 37(5): 235-248.

- Ridgen, D. (2009). *From protein structure to function with bioinformatics*. Springer, Liverpool, UK.
- Roberts, C.J. (2014). Therapeutic Protein Aggregation: Mechanism, Design, and Control. *Trends Biotechnol* 32(7): 372-380.
- Sassi, P., Perticaroli, S., Comez, L., Lupi, L., Paolantoni, M., Fioretto, D. et Morresi, A. (2011). Reversible and irreversible denaturation processes in globular proteins: from collective to molecular spectroscopic analysis. *Raman spectroscopy*. Doi: 10.1002/jrs.3013.
- Shao, H., Xu, L. et Yan, Y. (2014). Biochemical characterization of carboxylesterase from the Archaeon *Pyrobaculum* sp. 1860 and a rational explanation of its substrate specificity and thermostability. *International journal of molecular sciences* 15: 16 885-16 910.
- Science Editorial (2005). So much to know. *Science* 309: 78-102.
- Tsai, et al. (2015). Electrostatics, structure prediction, and the energy landscapes for protein folding and binding. *Protein science* 0(0): 1-15.
- Tsoli, A.C., Papandreou, N.C., Iconomidou, V.A. et Hamodrakas, S.J. (2013) A consensus method for the prediction of aggregation-prone peptides in globular proteins. *Plos One* 8(1): e54175.
- Van den Berg, B. (2003) Extremophiles as a source for novel enzymes. *Curr Opin Microbiol* 6: 213-218.
- Voet, D., Voet, J.G. et Pratt C. W. (2008) *Fundamentals of Biochemistry*, 3rd edition.
- Vetriani, C., Maeder, D.L., Tolliday, N., Yip, K.S., Stillman, T.J. et al. (1998) Protein stability above 100 degrees C: a key role for ionic interactions. *Proc Natl Acad Sci USA* 95: 12300-12305.
- Yu S., Yu, S., Han, W., Wang, H., Zheng, B. et Feng, Y. (2010). A novel thermophilic lipase from *Fervidobacterium nodosum* Rt17-B1 representing a new subfamily of bacterial lipases. *J Mol Catal B Enzym* 66: 81-89.